

JUAN RAFAEL NUÑEZ

**SEGMENTAÇÃO AUTOMÁTICA E
CLASSIFICAÇÃO EM VÍDEOS ESPORTIVOS**

Dissertação apresentada ao Programa de Pós-Graduação em Informática Aplicada da Pontifícia Universidade Católica do Paraná como requisito parcial para obtenção do título de Mestre em Informática Aplicada.

CURITIBA

2006

JUAN RAFAEL NUÑEZ

**SEGMENTAÇÃO AUTOMÁTICA E
CLASSIFICAÇÃO EM VÍDEOS ESPORTIVOS**

Dissertação apresentada ao Programa de Pós-Graduação em Informática Aplicada da Pontifícia Universidade Católica do Paraná como requisito parcial para obtenção do título de Mestre em Informática Aplicada.

Área de Concentração: Processamento Digital de Imagens

Orientador: Prof. Dr. Jaques Facon

CURITIBA

2006

N972s Nuñez, Juan Rafael
2006 Segmentação automática e classificação em vídeos esportivos / Juan Rafael Nuñez ; orientador, Jaques Facon. – 2006.
xviii, 179 f. : il. ; 30 cm

Dissertação (mestrado) – Pontifícia Universidade Católica do Paraná, Curitiba, 2006
Inclui bibliografia

1. Processamento de imagens – Técnicas digitais. 2. Vídeo digital.
3. Futebol. 4. Sistemas de indicação visual tridimensional. 5. Sistemas multimídias. I. Facon. JaCques. II. Pontifícia Universidade Católica do Paraná. Programa de Pós-Graduação em Informática Aplicada. III. Título.

CDD-21.ed. 621.367
006.6

Dedico

À Deus, à minha família e à Ciência.

Agradecimentos

Em especial à Perkons, que me incentivou e me proporcionou os meios necessários para a continuação de minha história acadêmica e profissional.

Aos meus pais, pelo estímulo à busca do saber e pela sólida formação que me foi dada, nunca medindo esforços para que eu tivesse sempre o melhor.

À minha esposa, Camila, pelo carinho e atenção dedicados.

Ao professor Dr. Jaques Facon pelo acompanhamento e orientação.

Agradeço de todo o coração a Deus, que me acompanhou me dando forças e me cercado de pessoas que me ajudaram em todo o tempo.

Sumário

Agradecimentos	v
Sumário.....	vi
Lista de Figuras	ix
Lista de Tabelas.....	xiii
Lista de Abreviaturas.....	xv
Resumo	xvii
Abstract.....	xviii

Capítulo 1

Introdução	1
1.1. Objetivos.....	3
1.2. Desafios	4
1.3. Contribuições.....	5

Capítulo 2

Espaços de Cores.....	6
2.1. Formato de Cor RGB.....	7
2.2. Formato de Cor HSV	8
2.3. Formato de Cor CMYK.....	10
2.4. Formato de Cor YCbCr	11

Capítulo 3

Classificação	13
3.1. Introdução.....	13

3.2.	Classificação Supervisionada vs Não Supervisionada	16
3.3.	Classificação Exclusiva vs Não Exclusiva	16
3.4.	Algoritmos e Processos de Classificação	17
3.5.	Considerações	25

Capítulo 4

Morfologia Matemática	26
4.1. Introdução	26
4.2. Operações básicas.....	27
4.3. Erosão	29
4.4. Dilatação	30
4.5. Propriedades da erosão e da dilatação	31
4.6. Operações geodésicas e reconstrução.....	32
4.7. Considerações.....	34

Capítulo 5

Segmentação de imagens	35
------------------------------	----

Capítulo 6

Imagem e Vídeo.....	38
6.1. Qualidade.....	38
6.2. Compressão	40
6.3. CODECs de vídeo	42

Capítulo 7

Estado da Arte	45
----------------------	----

Capítulo 8

Metodologia Proposta.....	54
8.1. Conversão RGB -> HSV	56
8.2. Estatística sobre o canal H.....	56
8.3. Limiarização (binarização).....	60
8.4. Morfologia Matemática e Ferramentas Matemáticas	61
8.5. Agrupamento	65

Capítulo 9

Experimentos e Discussão de Resultados.....	76
9.1. Considerações iniciais	76
9.2. Vídeos utilizados	76
9.3. Qualidade.....	78
9.4. Resultados.....	84
9.5. Tempo de Processamento	91
9.6. Problemas encontrados	95

Capítulo 10

Conclusão e Trabalhos Futuros	102
-------------------------------------	-----

Referências Bibliográficas.....	104
---------------------------------	-----

Apêndice A

Planilhas de Resultados	111
-------------------------------	-----

Apêndice B

Exemplos de imagens segmentadas e classificadas.....	140
--	-----

Apêndice C

Índices de qualidade para novos pesos	173
---	-----

Lista de Figuras

Figura 2.1 – Espectro de cores visíveis	7
Figura 2.2 – Sistema de cores RGB.....	7
Figura 2.3 – Sistema de cores HSV.....	9
Figura 2.4 – Representação do espaço de cores HSV	10
Figura 2.5 – Sistema de cores CMYK.....	11
Figura 2.6 – Sistema de cores YCbCr	12
Figura 4.1 – Exemplo de elementos estruturantes 3x3 e 5x5.....	28
Figura 4.2 – Distância geodésica.....	32
Figura 4.3 – Imagem Σ contendo partículas parcialmente incluídas (a); dilatação geodésica de $\partial \Sigma$ em X (b); subconjunto das partículas totalmente incluídas em Σ (c).	33
Figura 6.1 – Resultado da comparação MOS.....	41
Figura 8.1 – (a) Imagem de partida de futebol, (b) Imagem com cores detectadas como gramado em preto	59
Figura 8.2 – Histograma do canal H da Figura 8.1a.....	59
Figura 8.3 – (S_1) acumulação de histograma, (D_1) derivada de S_1 , (Mo_1) primeiro ponto onde D_1 se torna zero depois do pico	60
Figura 8.4 – (S_2) acumulação de histograma, (D_2) derivada de S_2 , (Mo_2) primeiro ponto onde D_2 se torna zero depois do pico	60
Figura 8.5 – Espaço de cor HSV de forma horizontal (a) e de forma circular (b)	61
Figura 8.6 – Seqüência de operações morfológicas	64
Figura 8.7 – Influência das regiões adjacentes na imagem: (a) Imagem Original, (b) Linha ampliada.	68
Figura 8.8 – Histograma da região segmentada 1	70
Figura 8.9 – Histograma da região segmentada 2	70
Figura 8.10 – Histograma da região segmentada 3	71
Figura 8.11 – Histograma da região segmentada 4	71
Figura 8.12 – Histograma da região segmentada 5	72
Figura 8.13 – Histograma da região segmentada 6	72

Figura 8.14 – Histograma da região segmentada 7	73
Figura 8.15 – Histograma da região segmentada 8	73
Figura 8.16 – Histograma de região segmentada	74
Figura 9.1 – (a) Exemplo de marcação manual de imagem, (b) Imagem Original	81
Figura 9.2 – Exemplo de marcação correta (branco) e incorreta (vermelho).....	82
Figura 9.3 – Oclusão/fusão de jogador.....	83
Figura 9.4 – (a) é a imagem original, (b) a imagem de marcação manual, (c) a imagem limiarizada, e (d) representa a diferença entre a marcação manual e a automática.....	84
Figura 9.5 – Exemplo de imagem já classificada	86
Figura 9.6 – Exemplo de imagem já classificada	88
Figura 9.7 – Gráfico de Tempo de Processamento x Pixels.....	91
Figura 9.8 - Gráfico de Tempo de Processamento (s) por Pixel x Pixel	92
Figura 9.9 – Gráfico de Tempo de Processamento por Etapa	93
Figura 9.10 – Exemplos de imagens não aproveitadas.....	94
Figura 9.11 – Diferença de tonalidade entre dois pontos na imagem, (a) e (b) são os pontos originais e (c) e (d) seus correspondentes de cor pura (saturação e valor em 100%).	95
Figura 9.13 – Detalhes do espalhamento das cores	97
Figura 9.14 – Detalhes do espalhamento das cores	97
Figura 9.15 – Diferença das tonalidades de cor de dois jogadores extraídos de uma mesma imagem	98
Figura 9.16 – Problema de imagem borrada: (a) e (b) Imagens captadas de um jogo, (c) e (d) Zoom dos jogadores marcados em (a) e (b).	99
Figura 9.17 – Confusões causadas por ângulo de visão	99
Figura 9.18 – Diferença da camisa dos jogadores em um mesmo quadro	100
Figura 9.19 – Região externa (a) classificada como campo, e seguranças classificados como jogadores (b)	101
Figura A.1– Exemplo de Classificação	112
Figura A.2 – Exemplo de Classificação	114
Figura A.3 – Exemplo de Classificação	116
Figura A.4 – Exemplo de Classificação	118
Figura A.5 – Exemplo de Classificação	120
Figura A.6 – Exemplo de Classificação	122
Figura A.7 – Exemplo de Classificação	124

Figura A.8 – Exemplo de Classificação	126
Figura A.9 – Exemplo de Classificação	128
Figura A.10 – Exemplo de Classificação	130
Figura A.11 – Exemplo de Classificação	132
Figura A.12 – Exemplo de Classificação	134
Figura A.13 – Exemplo de Classificação	136
Figura A.14 – Exemplo de Classificação	138
Figura B.1 – Exemplo.....	141
Figura B.2 – Exemplo.....	142
Figura B.3 – Exemplo.....	143
Figura B.4 – Exemplo.....	144
Figura B.5 – Exemplo.....	145
Figura B.6 – Exemplo.....	146
Figura B.7 – Exemplo.....	147
Figura B.8 – Exemplo.....	148
Figura B.9 – Exemplo.....	149
Figura B.10 – Exemplo.....	150
Figura B.11 – Exemplo.....	151
Figura B.12 – Exemplo.....	152
Figura B.13 – Exemplo.....	153
Figura B.14 – Exemplo.....	154
Figura B.15 – Exemplo.....	155
Figura B.16 – Exemplo.....	156
Figura B.17 – Exemplo.....	157
Figura B.18 – Exemplo.....	158
Figura B.19 – Exemplo.....	159
Figura B.20 – Exemplo.....	160
Figura B.21 – Exemplo.....	161
Figura B.22 – Exemplo.....	162
Figura B.23 – Exemplo.....	163
Figura B.24 – Exemplo.....	164
Figura B.25 – Exemplo.....	165
Figura B.26 – Exemplo.....	166
Figura B.27 – Exemplo.....	167

Figura B.28 – Exemplo.....	168
Figura B.29 – Exemplo.....	169
Figura B.30 – Exemplo.....	170
Figura B.31 – Exemplo.....	171
Figura B.32 – Exemplo.....	172
Figura C.1 – Índices de qualidade	174
Figura C.2 – Índices de qualidade	175
Figura C.3 – Índices de qualidade	176
Figura C.4 – Índices de qualidade	177
Figura C.5 – Índices de qualidade	178

Lista de Tabelas

Tabela 6.1 – CODECs utilizados.....	40
Tabela 6.2 – Colocação na avaliação de qualidade dos especialistas, da melhor para a pior qualidade.	41
Tabela 6.3 – Correlação entre MOS e outras técnicas objetivas, onde 1 reflete 100% de similaridade.	42
Tabela 8.1 – Detecção de campo manual/automática.....	58
Tabela 9.1 – Descrição dos vídeos usados	77
Tabela 9.2 – Legenda dos nomes dos vídeos.....	78
Tabela 9.3 – Critérios de medição de qualidade.....	79
Tabela 9.4 – Descrição da marcação manual das imagens.....	80
Tabela 9.5 – Exemplo de tabela de dados	86
Tabela 9.6 – Exemplo de tabela de resultados	87
Tabela 9.7 – Exemplo de tabela de dados	89
Tabela 9.8 – Exemplo de tabela de resultados	89
Tabela 9.9 – Resultado da análise dos vídeos	91
Tabela 9.10 - Tempo médio de processamento para diferentes resoluções de imagem.	93
Tabela A.1 – Tabela de dados	112
Tabela A.2 – Tabela de Resultados	113
Tabela A.3 – Tabela de dados	114
Tabela A.4 – Tabela de Resultados	115
Tabela A.5 – Tabela de dados	116
Tabela A.6 – Tabela de Resultados	117
Tabela A.7 – Tabela de dados	118
Tabela A.8 – Tabela de Resultados	119
Tabela A.9 – Tabela de dados	120
Tabela A.10 – Tabela de Resultados	121
Tabela A.11 – Tabela de dados	122

Tabela A.12 – Tabela de Resultados	123
Tabela A.13 – Tabela de dados	124
Tabela A.14 – Tabela de Resultados	125
Tabela A.15 – Tabela de dados	126
Tabela A.16 – Tabela de Resultados	127
Tabela A.17 – Tabela de dados	128
Tabela A.18 – Tabela de Resultados	129
Tabela A.19 – Tabela de dados	130
Tabela A.20 – Tabela de Resultados	131
Tabela A.21 – Tabela de dados	132
Tabela A.22 – Tabela de Resultados	133
Tabela A.23 – Tabela de dados	134
Tabela A.24 – Tabela de Resultados	135
Tabela A.25 – Tabela de dados	136
Tabela A.26 – Tabela de Resultados	137
Tabela A.27 – Tabela de dados	138
Tabela A.28 – Tabela de Resultados	139
Tabela C.1 – Tabela de pesos (pesos originais).....	174
Tabela C.2 – Tabela de dados por vídeo	174
Tabela C.3 – Tabela de pesos	175
Tabela C.4 – Tabela de dados por vídeo	175
Tabela C.5 – Tabela de pesos	176
Tabela C.6 – Tabela de dados por vídeo	176
Tabela C.7 – Tabela de pesos	177
Tabela C.8 – Tabela de dados por vídeo	177
Tabela C.9 – Tabela de pesos	178
Tabela C.10 – Tabela de dados por vídeo	178

Lista de Abreviaturas

- ADJ – Percentual de acerto na classificação de pixels de regiões adjacentes
- CIA – Total de pixels de campo (Imagem Manual)
- CIE – Commission Internationale de 'Eclairage
- CIM – Total de pixels de campo classificados corretamente
- CMPO – Percentual de acerto na classificação de pixels de campo
- CMYK – Cyan, Magenta, Yellow, Black (Ciano, Magenta, Amarelo, Preto)
- CODEC – Codificador/Decodificador de vídeo.
- DTW – Dynamic time warping
- HDTV – Televisão de alta definição (High-definition television)
- HSV – Hue, Saturation, Value (Tonalidade ou cor, Saturação, Valor)
- IQ1 – Índice de qualidade
- IQ2 – Índice de qualidade (Desconsiderando jogadores colados nas bordas)
- JCA – Número total de jogadores colados nas extremidades do campo
- JCC – Número total de jogadores classificados corretamente
- JCC1 – Jogadores classificados corretamente
- JCC2 – Jogadores classificados corretamente (Desconsiderando jogadores colados nas bordas)
- JNS – Número total de jogadores não segmentados
- JPG – Joint Photographic Experts Group
- JSC1 – Jogadores segmentados corretamente
- JSC2 – Jogadores segmentados corretamente (Desconsiderando jogadores colados nas bordas)
- MOS – Mean Opinion Score
- MP3 – MPEG-1 Audio Layer 3
- MPEG – Moving Picture Experts Group
- PSNR – Peak Signal-To-Noise Ratio (Relação Sinal Ruído)
- RAIA – Total de pixels de região adjacente campo (Imagem Manual)

RAIM – Total de pixels de região adjacente classificados corretamente

RCC – Regiões classificadas corretamente

RCCP – Regiões classificadas corretamente

RGB – Red, Green, Blue (Vermelho, Verde, Azul)

SSIM – Structural SIMilarity

TJ – Número total de jogadores visíveis na imagem

TRS – Total de regiões segmentadas

VQM – Video Quality Metric

vs – Versus

Resumo

A extração de características e recuperação de informações visuais de imagens digitais é uma área que cresce juntamente com as tecnologias de geração e armazenamento de informações digitais. A demanda por melhores soluções de gestão e manipulação deste conteúdo é crescente. Propõe-se aqui, prover subsídios para a criação de novas aplicações e a melhoria de outras já existentes através de uma nova técnica de segmentação automática de vídeos esportivos e classificação de elementos semânticos. A introdução deste trabalho traz uma lista de aplicações que podem ser beneficiadas pelas técnicas descritas. O desenvolvimento consiste essencialmente de uma nova técnica de segmentação automática, uma nova técnica de classificação de elementos, e uma nova técnica para mensurar o resultado deste tipo de processo. Testes foram feitos com um total de 14 vídeos variados em termos de resolução, qualidade, condições de luminosidade, campo e uniformes dos jogadores. Os resultados alcançados são promissores, visto que foram consistentes nos diferentes jogos testados, atingindo as expectativas e em alguns casos superando-as. O ponto forte deste trabalho foi a detecção de região dominante, que mesmo sendo um processo automático e não paramétrico, chegou na maioria dos casos a aproximadamente 98%. No índice geral, métrica desenvolvida neste trabalho, e que considera de maneira ponderada a segmentação e a classificação, a maioria dos vídeos analisados ficou com resultados acima de 80%.

Palavras-Chave: 1. Segmentação 2. Classificação 3. Automática 4. Futebol.

Abstract

The feature extraction and recovery of visual information of digital images is an area that grows with the technologies of generation and storage of digital information. The demand for better solutions of management and manipulation of this content is increasing. We propose to provide subsidies for the creation of new applications and the improvement of existing ones a new technique of automatic segmentation of sports videos and classification of semantic elements. The introduction of this work brings a list of applications that can be benefited by the described techniques. The development consists essentially of a new technique of automatic segmentation, a new technique of classification of elements, and a new technique to measure the results. Tests have been made with a total of 14 videos, different in terms of resolution, quality, and conditions of luminosity, field and clothes of the players. The reached results are promising, therefore they had been consistent in the different tested games, reaching the expectations and in some cases surpassing them. The strong point of this work was the detection of dominant region, that even being a process automatic and without parameters, reached in the majority of the cases 98%. In the general index, metric developed in this work, which considers in a weighed way the segmentation and the classification, the majority of the analyzed videos reached results above of 80%.

Keywords: 1. Segmentation 2. Automatic 3. Classification 4. Soccer.

Capítulo 1

Introdução

A extração de características e recuperação de informações visuais de imagens digitais é uma área que cresce juntamente com as tecnologias de geração e armazenamento de informações digitais.

Existem inúmeras aplicações que necessitam melhores soluções e tecnologias para o processamento digital de imagens, como:

- a. Sumarização automática de partidas esportivas;
 - Gerar resumos da partida apenas com o conteúdo mais importante, eliminando os intervalos, lances parados em cobranças de falta/escanteios/laterais, entrevistas, comerciais, etc.
 - Dar ao usuário mais recursos para escolher o que ele realmente quer ver.
- b. Localização de cenas baseadas em consulta de alto nível;
 - Localizar cenas de conteúdo específico na partida, como os lances de ataque, as cobranças de bola parada, escanteios, faltas, gols, etc.
- c. Análise automática do jogo;
 - Analisar e calcular automaticamente o de tempo de domínio de bola, quantidade de faltas, tiros de meta, escanteios, pênaltis, etc.
- d. Detecção de cortes, comerciais, narrações etc.
 - Suprimir ou aumentar o nível de compressão destes elementos.
- e. Detecção de ângulo e zoom de câmera;
 - Identificar tomadas em zoom, longa distância, saber qual a parte do campo está sendo focada.
- f. Reconstrução tridimensional das partidas;

- É possível, usando várias câmeras, reconstruir uma partida tridimensionalmente, para análise tática, simulações do ponto de vista dos jogadores, juiz, etc.
- g. Escolha automática da melhor cena no momento da geração dos vídeos;
- Com múltiplas câmeras em campo, é possível tempo real saber quais as câmeras que estão captando o jogo num ângulo melhor.
- h. Adequação para dispositivos portáteis;
- Está se tornando cada vez mais popular a transmissão de eventos esportivos para dispositivos como celulares, *palmtops*, etc. A visualização neste tipo de dispositivo ainda é muito limitada, e o conteúdo a ser transmitido é diferente do transmitido para uma televisão. Como a produção de um conteúdo específico se torna muito cara, a transformação automática do conteúdo existente é a alternativa mais interessante. Para isso é possível enquadrar a imagem na área mais importante antes de redimensioná-la.
- i. Assistência aos árbitros;
- Alguns campeonatos de tênis já usam sistemas de vídeo para tirar dúvidas sobre posicionamento de bolas muito próximas às linhas do campo. É possível utilizar sistemas de visão computacional para auxiliar em tarefas mais complexas, nas quais não se tem uma linha fixa, como por exemplo, em lances de impedimento no futebol. No futuro este tipo de solução será mais utilizada, pois decisões erradas costumam envolver cifras muito altas.
 - A atual crise de corrupção nas arbitragens esportivas em vários países está criando o cenário para a adoção de mecanismos eletrônicos mais sofisticados, para o auxílio instantâneo dos árbitros nas partidas.
- j. Auxílio na apresentação de jogos;
- A apresentação de esportes pode ser melhorada com várias técnicas, como o uso de ilustrações no vídeo original, como a superposição de linhas descrevendo trajetórias, do número do jogador que pode estar obstruído, linhas mostrando a distância entre jogadores ou da bola, ou até mesmo comerciais em regiões específicas.
- k. Melhoria da exibição;
- O placar e eventuais propagandas ou efeitos visuais muitas vezes podem ocultar parcialmente um lance. A exibição pode ser aprimorada se estes efeitos visuais sumirem quando houver um lance na mesma região do vídeo.

Separar automaticamente as informações semânticas em vídeos é um dos grandes desafios da visão computacional, e ao mesmo tempo é um requisito básico para a maior parte das aplicações citadas. Nos casos onde se tem uma câmera móvel, como nas filmagens esportivas, a tarefa é ainda mais complexa, pois grande parte das técnicas já desenvolvidas não são eficazes, pois dependem de um plano de fundo estático.

O número crescente de canais de televisão e eventos esportivos, somado ao tempo cada vez menor do público para acompanhar as partidas de interesse, está demandando técnicas mais eficientes de captura, seleção e armazenamento deste tipo de conteúdo digital. Para exemplificar, um cliente de tv por assinatura, que tem 2 horas ao final do dia para ver televisão, está pagando pela transmissão dos 4 jogos que passaram ao vivo durante a tarde, mas não poderá ver nenhum deles. Poderia se o aparelho de tv fosse capaz de gravar as partidas, selecionar as tomadas mais relevantes, e gerar um compacto de cada uma delas. Com este tipo de tecnologia, o cliente poderia nestas 2 horas ver não só um compacto com as 4 partidas, como também o jornal ou outro programa.

1.1. Objetivos

O objetivo deste trabalho é desenvolver uma técnica robusta de processamento para a segmentação e classificação em vídeos esportivos.

Esta tarefa será realizada quadro a quadro, tornando-se independente de informações de quadros anteriores, e podendo ser aplicada tanto em seqüências de vídeo como em imagens estáticas. O processo deverá ser completamente automático, não paramétrico, e compatível com a grande variedade de resoluções, e qualidades de vídeos mais comuns.

O esporte escolhido para a elaboração dos testes foi o futebol, pela diversidade de vídeos disponíveis, tanto em termos de qualidade como resolução, tipos de gramado, estádios, sombras projetadas em campo, variedade de camisas dos times, e especialmente, a popularidade do esporte.

Os algoritmos desenvolvidos neste trabalho devem ser testados com vídeos de diferentes resoluções e qualidades, de modo que sua robustez e eficiência possam ser mensuradas e comprovadas.

1.2. Desafios

Um quadro de vídeo capturado de uma transmissão de TV, e uma foto tirada com uma máquina fotográfica digital de última geração trazem uma diferença muito grande em termos de resolução e qualidade. As transmissões de TV trazem uma imagem de baixa resolução, com perda de qualidade pela transmissão, compressão dos dados, ruído da digitalizadora etc. Esta imagem de péssima qualidade passa despercebida, pois normalmente é vista por apenas aproximadamente 0,03 segundos (na taxa de amostragem de televisão, que é de aproximadamente 30 quadros por segundo). Ao congelar uma imagem e analisá-la detalhadamente, é possível perceber que seus contornos, contraste, e cores já não são nem próximos aos de uma fotografia. Elaborar um algoritmo capaz de segmentar uma imagem de vídeo deste tipo é na verdade trabalhar com o que existe de pior em termos de qualidade de imagem. É preciso ter em mente que nas imagens esportivas, na maior parte do tempo os elementos vão aparecer borrados, desfocados, sem contornos e com suas cores e formas comprometidas pelos fatores já mencionados. Os trinta quadros por segundo iludem a vista humana, mas quando são processados individualmente se mostram um sério problema.

É importante mencionar que apesar das heurísticas e os parâmetros normalmente facilitarem muito o trabalho, poupando tempo, e simplificando os algoritmos, neste trabalho, um dos desafios é evitá-los por completo, para que os conceitos aqui criados possam ser aproveitados por outras aplicações e estudos.

Mesmo nas imagens de mais alta definição, como as de qualidade de DVD ou TV digital, os jogadores são representados por algumas poucas dezenas de pixels. Separar estas pequenas porções de pixels, e agrupá-las de maneira semântica também é algo longe de ser elementar, requerendo a aplicação de conhecimentos matemáticos, estatísticos, óticos, físicos e computacionais.

1.3. Contribuições

Algumas das contribuições deste trabalho de pesquisa para o desenvolvimento da área são:

- Um novo algoritmo para identificação automática dos limiares na segmentação de região predominante;
- Uma nova técnica não supervisionada de agrupamento de regiões baseada em cor, funcional inclusive para pequenas regiões;
- Uma nova métrica para mensurar a qualidade do resultado da segmentação de imagens esportivas;
- Um banco de imagens variadas de futebol, bem como os resultados da análise pelos algoritmos propostos;
- Uma plataforma de processamento automática e genérica, que pode ser especializada para várias aplicações, tanto para dar continuidade a este trabalho como para contribuir com pesquisas relacionadas ao tema do presente projeto.

Capítulo 2

Espaços de Cores

A cor é um aspecto físico da natureza. A cor de um material é determinada pelos comprimentos de onda dos raios luminosos que as suas moléculas constituintes refletem. Um objeto terá determinada cor se não absorver justamente os raios correspondentes à frequência daquela cor. Assim, um objeto que é vermelho, absorve todos os raios de luz, exceto os que têm o comprimento de onda relativo à frequência do vermelho.

A cor é relacionada com os diferentes comprimentos de onda do espectro eletromagnético, e é percebida pelas pessoas na faixa da zona visível e por alguns animais através dos órgãos de visão, como uma sensação que nos permite diferenciar os objetos do espaço com maior precisão.

Considerando as cores como luz, a cor branca resulta da superposição de todas as cores, enquanto o preto é a ausência de luz. Uma luz branca pode ser decomposta em todas as cores (o espectro) por meio de um prisma. Na natureza, esta decomposição origina um arco-íris.

Logo, tudo o que se vê, reflete uma cor e absorve todas as outras. E dependendo das condições de iluminação a que um objeto está submetido, são vistos matizes diferentes da cor refletida e até mesmo outras cores que antes haviam sido absorvidas.

O que as cores (denominadas de luz visível), representam é uma parcela diminuta do espectro eletromagnético, se estendendo de 400 nm a 700 nm aproximadamente, conforme pode ser visto na figura 2.1. É aceitado internacionalmente desde 1931 os valores 435,8 nm, 546,1 nm e 700 nm como os que representam espectralmente as três cores primárias aditivas, azul, verde e vermelho respectivamente, de acordo com a CIE (*“Commission Internationale de ‘Eclairage”*).

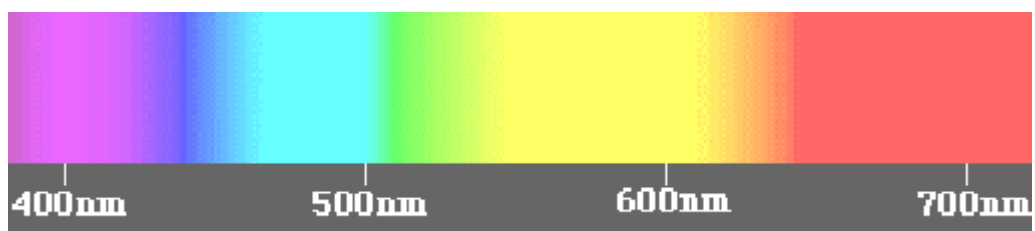


Figura 2.1 – Espectro de cores visíveis

2.1. Formato de Cor RGB

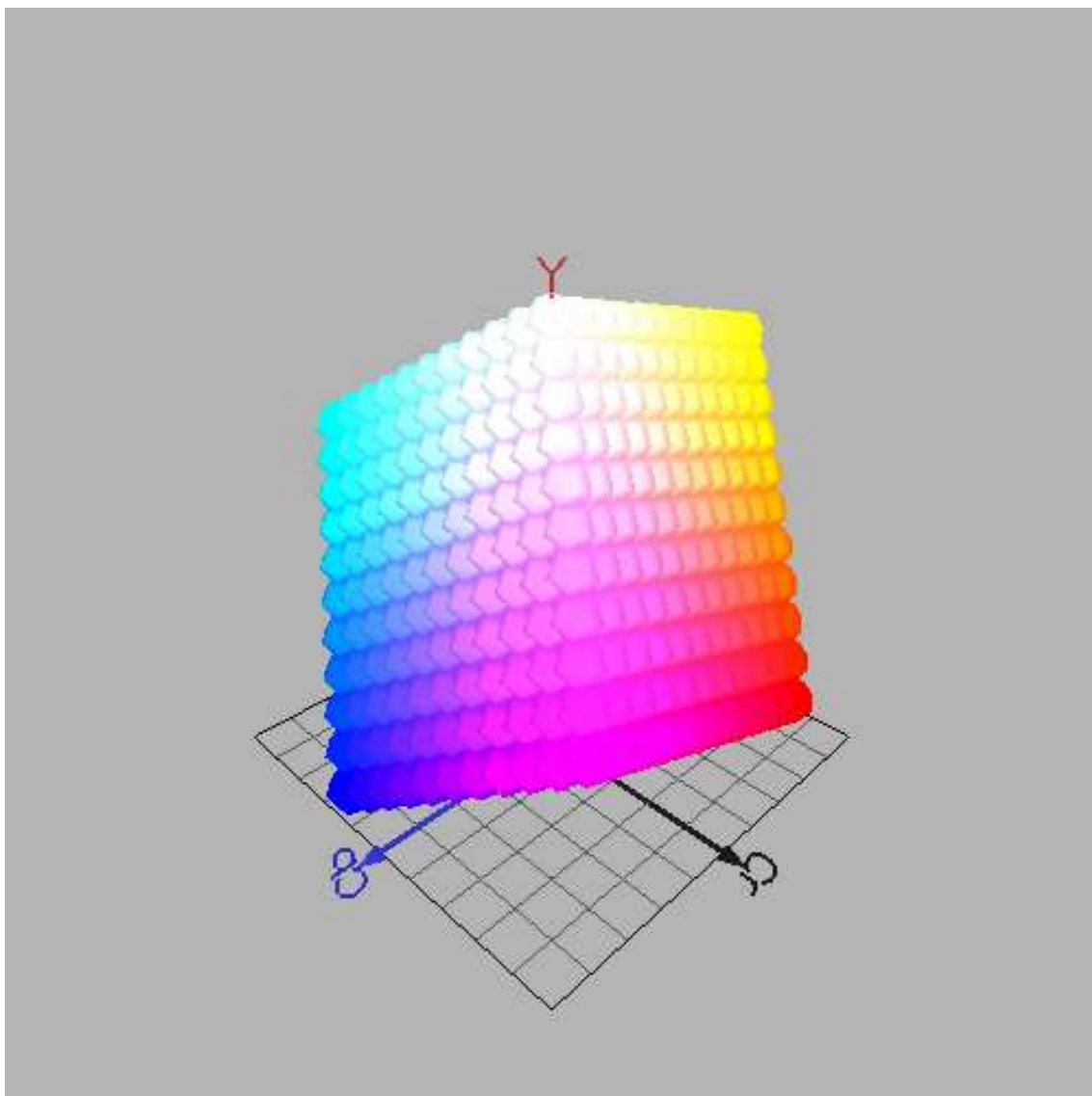


Figura 2.2 – Sistema de cores RGB

O formato de cor RGB não é capaz de representar todas as cores perceptíveis por um humano, mas é uma boa aproximação. O principal problema do sistema RGB é o fato de não permitir a impressão das cores que vemos na tela, os sistemas de impressão atuais (impressoras, *offset*, etc.) trabalham com o modo de cor CMYK.

As cores geradas no monitor de vídeo do computador são obtidas por um processo aditivo de três cores primárias: vermelho, verde e azul (*red, green, blue*). As primárias RGB são adequadas para gerar uma boa fração das cores visíveis porque o olho humano possui três tipos de detectores (cones) independentes, com respostas máximas próximas de R, G e B. Quando os três detectores são excitados com diferentes intensidades, ocorre o processo cognitivo de visão colorida. O processo é denominado aditivo porque decorre da adição das excitações dos três tipos de cones. A cor é então "formada" no olho.

As regras básicas de adição de cores são:

- Seja a cor $C = xR + yG + zB$
- $x = y = z \rightarrow$ tons de cinza de preto ($x=0$) e branco ($x=1$);
- $x = y, z=0 \rightarrow$ amarelo;
- $y=z, x=0 \rightarrow$ ciano;
- $x=z, y=0 \rightarrow$ magenta.

2.2. Formato de Cor HSV

HSV é a sigla para: *hue*, *saturation* e *value*, que traduzido seria: matiz, saturação e valor. É um formato de cor que separa a cromaticidade da luminância.

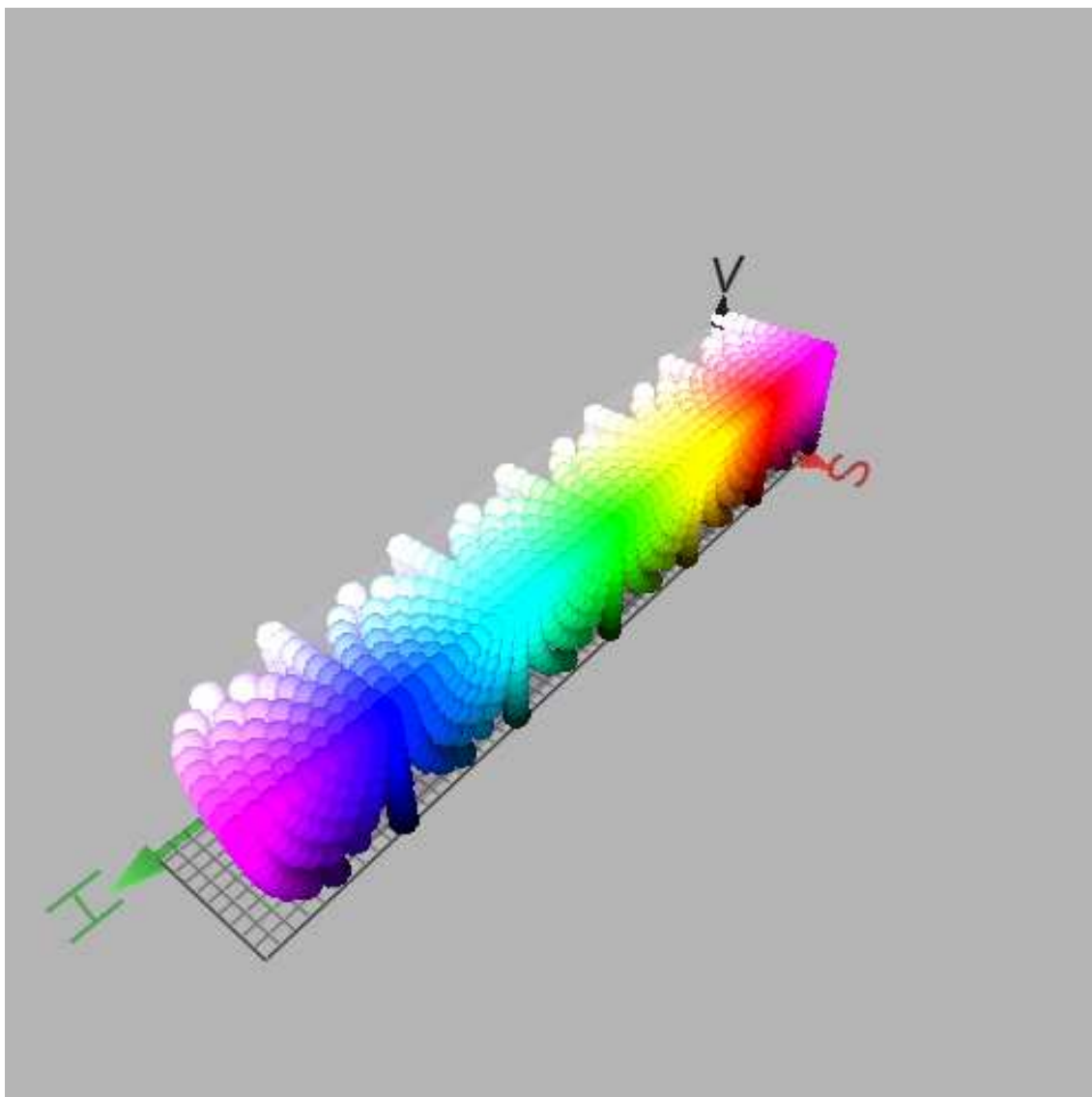


Figura 2.3 – Sistema de cores HSV

- **Matiz** é a cor pura dominante percebida pelo observador (vermelho, amarelo, azul, etc.). Por exemplo, se um corpo reflete maior parte de vermelho e menos de verde, é matiz alaranjado.
- **Saturação** é o quanto a cor pura dominante (matiz) é diluída pela luz branca. O alaranjado tem as mesmas cores do marrom, mas aparece alaranjado porque é mais saturado.
- **Valor** é a quantidade de luz refletida (claro/escuro), ou a quantidade de luz branca na cor. Por isso dizemos que as cores brilhantes são mais luminosas que as cores pastel.

A conversão do elemento H do HSV a partir do RGB é dada pela equação 2.1.

$$H = \begin{cases} 60 \times \frac{G-B}{MAX-MIN} + 0, & \text{if } MAX = R \\ & \text{and } G \geq B \\ 60 \times \frac{G-B}{MAX-MIN} + 360, & \text{if } MAX = R \\ & \text{and } G < B \\ 60 \times \frac{B-R}{MAX-MIN} + 120, & \text{if } MAX = G \\ 60 \times \frac{R-G}{MAX-MIN} + 240, & \text{if } MAX = B \end{cases} \quad (2.1)$$

Onde MAX representa o valor máximo e MIN o valor mínimo

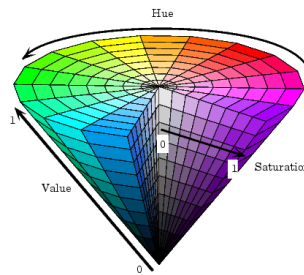


Figura 2.4 – Representação do espaço de cores HSV

2.3. Formato de Cor CMYK

A Sigla CMYK significa em inglês *Cyan*, *Magenta*, *Yellow* e *Black*, o que em português se traduz para ciano, magenta, amarelo, e preto. O preto surge como quarta cor porque nos sistemas de reprodução gráfica, impressão, a soma ou mistura dos pigmentos (da tinta) CMY não resulta em preto 100%, sendo necessário o acréscimo dessa quarta cor, o preto.

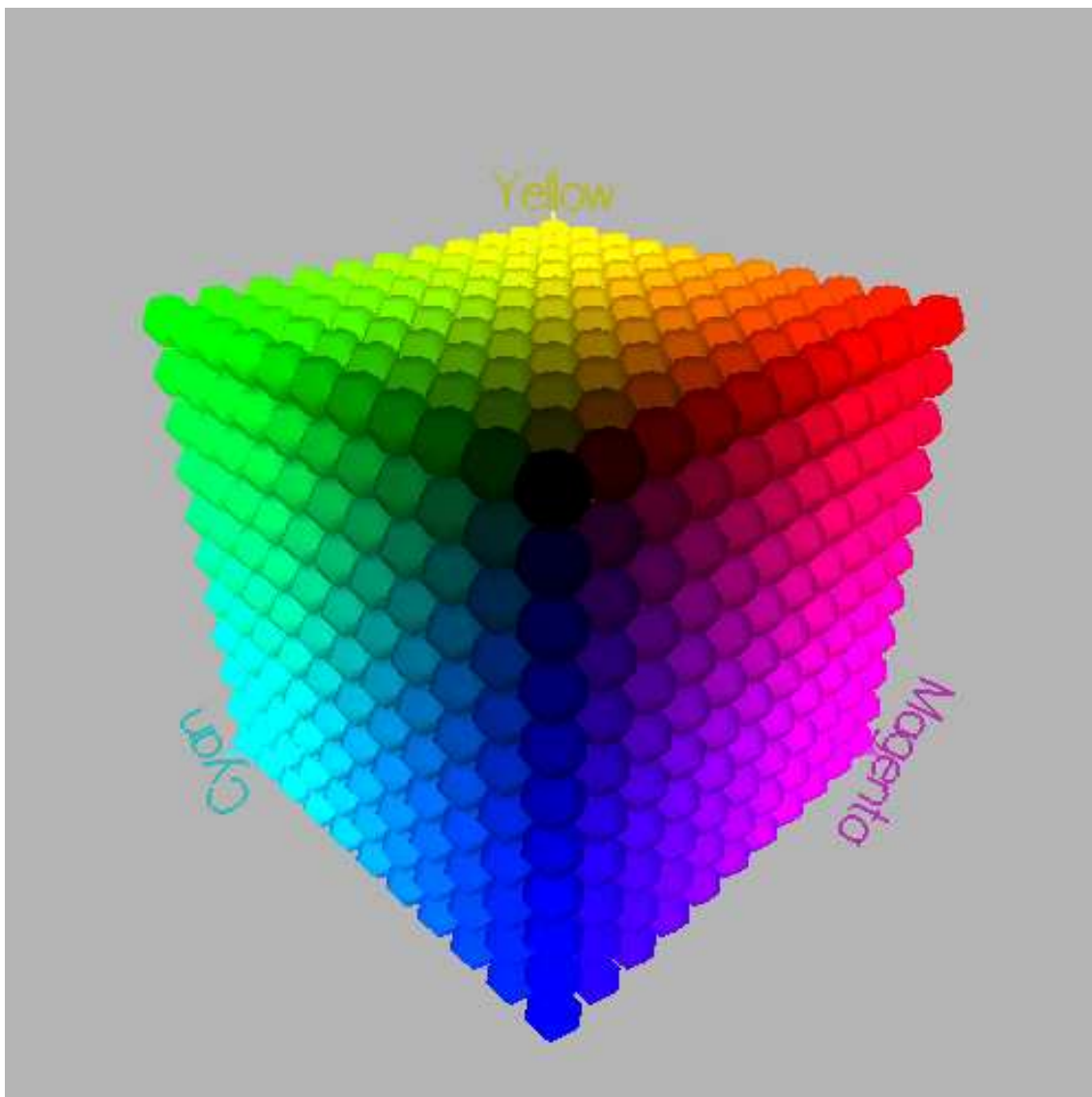


Figura 2.5 – Sistema de cores CMYK

2.4. Formato de Cor YCbCr

O modelo YCbCr é utilizado pelos canais de TV na Europa. A componente Y é chamada de "luminância" (*luminance*), e não possui nenhum efeito sobre a cor do pixel, mas apenas sobre sua luminosidade, regula o brilho da imagem. As componentes Cb e Cr são denominadas "crominância" (*chrominance*), e juntas definem a cor de cada pixel.

De forma simplificada, é possível resumir que as componentes Cb e Cr representam, respectivamente, a relação entre o azul e o vermelho em relação ao verde.

Estudos demonstram que a maior parte das informações mais sensíveis ao olho humano encontra-se na componente Y. As outras componentes Cb e Cr são menos

importantes para a formação da imagem, do ponto de vista de nossa percepção visual, logo aquela componente terá prioridade no descarte de informações para efeito de compressão.

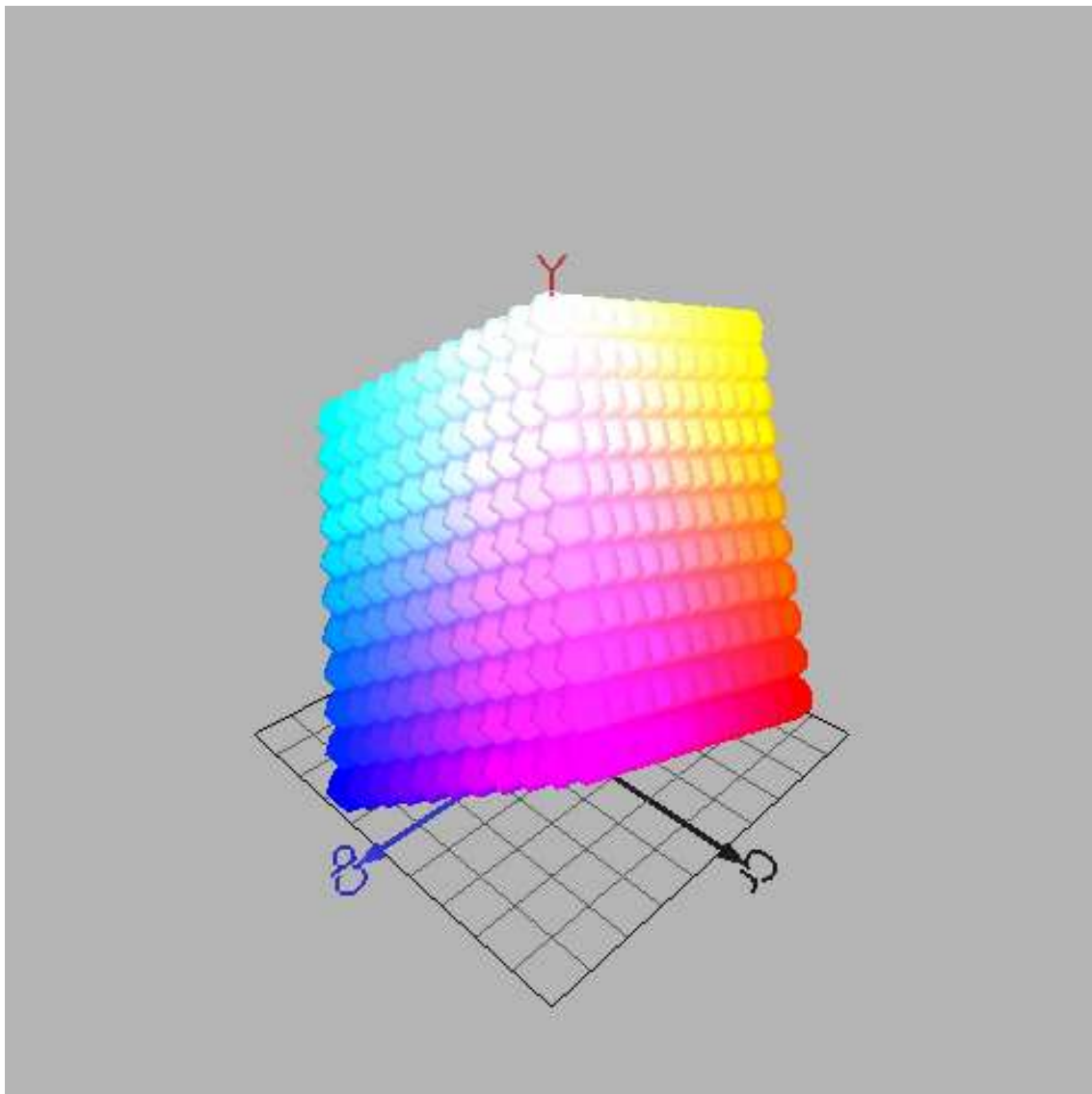


Figura 2.6 – Sistema de cores YCbCr

Capítulo 3

Classificação

3.1. Introdução

O processo de agrupamento é uma das mais primitivas capacidades mentais do ser humano, extremamente necessária para processar a incomensurável quantidade de informações adquiridas todos os dias, seja pela visão, audição, olfato, ou quaisquer outros sentidos em conjunto ou não. Lidar com cada objeto, som, cheiro, como uma entidade única seria impossível, logo, o ser humano cria categorias para classificar cada uma destas entidades. Como um exemplo, as pessoas tendem a classificar os animais em classes, como mamíferos, répteis, aves, etc. Isto permite inclusive algumas suposições, mesmo que nem sempre verdadeiras. Ao olhar para um animal com penas e asa, porém nunca visto antes, um ser humano normalmente tende a classificá-lo como uma ave, logo supõe que este animal desconhecido voa, e bota ovos.

Assim como para os seres humanos, o agrupamento é fundamental para análise estatística de informação, em praticamente todas as áreas da ciência, em especial cita-se a de visão computacional, recuperação de informações, reconhecimento de padrões, etc.

Por definição um grupo (*cluster*) é um conjunto de objetos de informação onde cada objeto é:

- Semelhante aos outros objetos do mesmo grupo;
- Não semelhante aos outros objetos dos outros grupos;

O objetivo do agrupamento então é a partir de um conjunto de dados sem rótulo, criar e organizar grupos, não importando se os dados são binários, categóricos, numéricos, um intervalo, ordinais, relacionais, textuais, espaciais, temporais, espaço-temporais, imagem, multimídia ou uma mistura desses tipos de dados. Os dados são chamados de estáticos se todas as suas características não se modificam ou a mudança é

muito pequena. A maior parte dos trabalhos de agrupamento usa dados estáticos, de modo que são classificados em cinco categorias principais:

- Métodos de particionamento;
- Métodos hierárquicos;
- Métodos baseados em densidade;
- Métodos baseados em grade;
- Métodos baseados em modelo;

Supondo um conjunto de “n” dados não-rotulados, um método de agrupamento constrói “k” grupos dos dados, onde cada grupo representa uma classe contendo pelo menos um objeto sendo $k \leq n$. A partição é chamada “*crisp*” se cada objeto pertence à exatamente uma classe, ou nebulosa se um objeto pode estar em mais de uma classe, em graus diferentes.

Dois métodos heurísticos para partições “*crisp*” bem conhecidos da literatura são o algoritmo *k-means* (ou k-médias), onde cada classe é representada pela média do valor dos objetos na classe e o algoritmo *k-medoids* (k-medianas), no qual cada classe é representada pelo objeto localizado mais ao centro do *cluster*. Outros métodos heurísticos de renome são o algoritmo *fuzzy c-means* e o algoritmo *fuzzy c-medoids*. Esses algoritmos heurísticos são eficazes para achar classes de formas esféricas e conjuntos pequenos ou médios de dados. Para localizar classes com formas não-esféricas, ou formas complexas, são usados algoritmos especiais, como o Gustafson-Kessel e algoritmos adaptativos de classificação nebulosa, podem também ser necessários métodos baseados em densidade para serem introduzidos na sequência.

Os métodos de classificação hierárquicos trabalham agrupando os dados a serem classificados em uma árvore de classes. Segundo [ANDRADE04] os dois tipos mais comuns de classificação hierárquica: a aglomerativa e a divisiva.

Métodos aglomerativos começam colocando cada objeto em sua classe e então funde as classes em classes cada vez maiores, até todos os objetos estarem classificados ou até que certos critérios de parada e o número de classes sejam satisfeitos. Métodos divisivos fazem exatamente o contrário. Um método de classificação hierárquico puro sofre na sua inabilidade de ajustar-se uma vez que a decisão de divisão ou união foi executada. Para melhorar a qualidade de classificação dos métodos hierárquicos, pode-se integrar a classificação hierárquica com outras técnicas de classificação.

A idéia geral dos métodos baseados em densidade é de continuar o crescimento de uma classe enquanto a densidade (número de objetos ou pontos de dados) na vizinhança não exceda algum critério. Ao invés de produzir uma classe explicitamente, é possível computar uma ordenação de classes aumentada para uma análise automática e interativa das classes. A ordenação contém informações que são equivalentes à classificação baseada em densidade obtida por um extenso número de parâmetros, desta forma eliminando a dificuldade de se escolher valores para os parâmetros.

Métodos baseados em grade quantizam o espaço de objetos em um número finito de células que formam uma estrutura de grade, onde todas as operações para classificação serão efetuadas. WANG usa vários níveis de células retangulares correspondendo a diferentes níveis de resolução. Informações estatísticas sobre os atributos de cada célula são pré-computadas e guardadas. Um processo de fila normalmente começa em um nível relativamente alto da estrutura hierárquica. Para cada célula na camada atual, o intervalo de confiança é computado refletindo a relevância da célula na fila dada. Células irrelevantes são removidas das próximas considerações. O processo da fila continua para o próximo nível mais baixo para as células relevantes, até a camada mais baixa ser atingida.

Métodos baseados em modelos assumem um modelo para cada classe tentando classificar cada dado no modelo mais parecido. Existem duas abordagens principais de métodos baseados em modelos: a abordagem estatística e a abordagem por redes neurais.

Assim, para dados estáticos, em séries temporais é preciso um algoritmo ou procedimento de classificação, sendo que a escolha do algoritmo depende tanto do tipo de dado disponível quanto do propósito particular da aplicação. Até onde vão os dados em séries temporais, distinções podem ser feitas, sendo o valor do dado discreto ou real, com amostras uniformes ou não, univariados ou multivariados, e se as séries de dados são de mesmo valor ou não. Dados cujas amostras não são uniformes devem ser convertidos em dados uniformes antes das operações de classificação serem executadas. Isso pode ser feito por um grande número de métodos, desde amostragem simples baseado no maior intervalo de amostras até uma abordagem mais sofisticada usando modelos e estimativas.

Vários algoritmos foram desenvolvidos para classificar tipos diferentes de dados em séries temporais. Deixando suas diferenças de lado, é justo dizer que o intuito de todas é tentar modificar os algoritmos existentes para classificação de dados estáticos de

uma maneira que os dados em séries temporais possam ser usados, ou converter os dados nas séries temporais em uma forma de dados estáticos para que os algoritmos já existentes de classificação em dados estáticos possam ser usados. A primeira abordagem geralmente usa diretamente dados não trabalhados em séries temporais, e por isso é chamada de abordagem baseada em dados não trabalhados (*raw-data-based approach*), e a maior modificação consiste em substituir a medição da distância/similaridade para dados estáticos em uma medição mais apropriada para séries temporais. Já a segunda abordagem converte os dados não trabalhados em um vetor de características de dimensão menor ou em um número de parâmetros de modelo, e então aplica um algoritmo convencional de classificação nos vetores de características extraídos ou nos parâmetros de modelo. Por este motivo são chamados de abordagens baseadas em características ou em modelo, respectivamente.

A maioria dos algoritmos/procedimentos de classificação é iterativa por natureza, e por isso, dependem de um critério para determinar quando uma boa classificação foi obtida, para poder parar o processo iterativo.

3.2. Classificação Supervisionada vs Não Supervisionada

A classificação supervisionada aplica somente uma medida de similaridade entre os pontos para se determinar a classe de um ponto. Esta classificação utiliza categorias já rotuladas, isto é, o número de grupos é conhecido. Uma classificação não supervisionada possui categorias ainda não rotuladas, isto é, os algoritmos não supervisionados precisam determinar o número de classes nos conjuntos de dados fornecidos, para então classificá-los.

3.3. Classificação Exclusiva vs Não Exclusiva

Uma classificação exclusiva é uma partição de um conjunto, onde cada ponto pertence exclusivamente a um único grupo. Uma classificação não exclusiva, ou sobreposta, pode assumir que um ponto pertença a vários grupos. Um agrupamento

nebuloso é um tipo de classificação não exclusiva, onde o ponto é associado a cada um dos grupos com um respectivo grau de pertinência.

3.4. Algoritmos e Processos de Classificação

Classificação com realocização

O processo de classificação com realocização tem segundo [LIAO05] os seguintes passos:

1. Começar com uma classificação inicial, denotada por C , já tendo o número “ k ” pré-estabelecido de classes.
2. Computar, para cada valor temporal, a matriz de dissimilaridade e guardar todas as matrizes resultantes computadas para todos os valores temporais para o cálculo de semelhanças na trajetória.
3. Procurar uma classificação C' , tal que C' seja melhor que C em termos da função critério generalizada Ward. A classificação C' é obtida a partir de C , realocando-se um membro de C_p para C_q , ou trocando dois membros entre C_p e C_q , onde $C_p, C_q \in C$, $p, q = 1, 2, \dots, k$ e $p \neq q$. Neste processo, se tal classificação não existir, deve-se interromper ou parar a realocização, ou então, colocar C' no lugar de C e repetir o passo 3.

Esse processo funciona apenas com séries temporais de igual comprimento, em função da distância entre duas séries temporais em algumas seções cruzadas (pontos de tempo onde uma série temporal não tem valor) não serem bem definidas.

Classificação hierárquica aglomerativa e divisiva

Classificação hierárquica é um método que funciona agrupando objetos de dado (nesse caso, séries temporais) em uma árvore de classes. Segundo [LIAO05] existem dois tipos de métodos de classificação hierárquica: aglomerativa e divisiva, dependendo da estratégia seguida, seja a *bottom-up* (começando mais de baixo, e indo para cima) ou *top-down* (começando de cima e indo para baixo). O método hierárquico aglomerativo é mais popular que o método divisivo. Ele começa posicionando cada objeto em sua própria classe e então funde essas classes atômicas (indivisíveis) em classes cada vez

maiores, até que todos os objetos estejam em uma única classe ou certas condições de parada sejam satisfeitas. O único algoritmo (completo) de ligação mede a similaridade entre duas classes como a similaridade do par de pontos de dados mais próximos pertencentes a classes diferentes. Então funde as duas classes com a menor distância, e repete o processo de união até que todos os objetos estejam eventualmente unidos em uma única classe. O algoritmo de variância mínima de Ward une duas classes que resultarão no menor crescimento no valor da variância da soma dos quadrados. A cada passo da classificação, todas as uniões possíveis entre duas classes são testadas. A variância da soma dos quadrados é computada para cada classe e o menor valor é selecionado.

Um método hierárquico aglomerativo normalmente sofre, em seu processo, de uma inabilidade de ajustar-se uma vez que a decisão de união entre duas classes foi executada. O mesmo é verdade para os métodos hierárquicos divisivos.

Classificação hierárquica não é restrita para a classificação de classes com séries temporais iguais, ou seja, também é aplicável em intervalos de tamanho desigual, se a medida de distância apropriada for usada para computar a distância ou a similaridade.

K-means e o fuzzy c-means

O *k-means* (também chamado de *c-means*) foi desenvolvido pela primeira vez há quase quatro décadas [MACQUEEN67]. A idéia principal deste método é a minimização de uma função objetiva, que normalmente é escolhida como sendo a distância total entre todas as características de seus respectivos centros de classe. Os dois passos principais do algoritmo *c-means*, e que denotam sua iteratividade são: a distribuição dos objetos entre as classes e a atualização dos centros de *cluster*.

O algoritmo alterna entre esses dois passos até que o valor da função objetiva não possa mais ser reduzido. Dadas n características $\{x_k | k=1, \dots, n\}$, o *c-means* determina c centros de classe $\{v_i | i=1, \dots, c\}$ minimizando a função de objetivo, definida pela equação 3.1.

$$MinJ_1(U, V) = \sum_{i=1}^c \sum_{k=1}^n \mu_{ik} \|x_k - v_i\|^2 \quad (3.1)$$

Onde: $\mu_{ik} \in \{0,1\} \forall i, k$, (2) $\sum_i \mu_{ik} = 1, \forall k$. $\| \cdot \|$ na equação acima normalmente é a medida de distância euclidiana. Contudo, outras medidas de distância podem ser usadas.

O processo de solução iterativa tem os seguintes passos:

1. Escolher c ($2 \leq c \leq n$) e ε (um número pequeno para parar o procedimento iterativo). Mude o contador $l = 0$ e os centros de classe iniciais V arbitrariamente.
2. Distribuir $x_k, \forall k$ para determinar U tal que J_1 seja minimizada. Isto é obtido normalmente realocando x_k para uma nova classe que está mais perto dele.
3. Atualizar os centros de classe V .

Se a mudança em V é menor que ε ; deve-se interromper o processo, caso contrário incrementar l e repetir os passos 2 e 3.

O algoritmo *c-means* pode ser estendido para trabalhar com partições nebulosas ao invés de trabalhar apenas com partições fixas, isto é demonstrado pela aplicação da fórmula 3.2:

$$MinJ_2(U, V) = \sum_{i=1}^c \sum_{k=1}^n (\mu_{ik})^2 \|x_k - v_i\|^2 \quad (3.2)$$

Observe que $U = [\mu_{ik}]$ nessa e nas próximas equações denota a matriz de uma c -partição nebulosa na qual, os confinamentos são (1) $\mu_{ik} \in [0,1] \forall i, k$, (2) $\sum_{i=1}^c \mu_{ik} = 1, \forall k$, e (3) $0 < \sum_{k=1}^n \mu_{ik} < n, \forall i$. Resumindo, cada x_k pode pertencer a mais de uma classe com cada índice de pertinência sendo um valor fracionário entre 0 e 1. É possível generalizar $J_2(U, V)$ para um número infinito de funções de objetivo, i.e. $J_m(U, V)$, onde $1 \leq m \leq \infty$. Assim, a nova fórmula para a mesma c -partição é mostrada na equação 3.3:

$$MinJ_m(U, V) = \sum_{i=1}^c \sum_{k=1}^n (\mu_{ik})^m \|x_k - v_i\|^2 \quad (3.3)$$

Ao se diferenciar a função de objetivo com respeito a v_i (para um U fixo) e com μ_{ik} (para um V fixo) respeitando as condições, se obtém as seguintes equações:

$$v_i = \frac{\sum_{k=1}^n (\mu_{ik})^m x_k}{\sum_{k=1}^n (\mu_{ik})^m}, i = 1, \dots, c. \quad (3.4)$$

$\mu_i = \frac{(1/\ x_k - v_i\ ^2)^{1/(m-1)}}{\sum_{j=1}^c (1/\ x_k - v_j\ ^2)^{1/(m-1)}}, i = 1, \dots, c; k = 1, \dots, n.$	(3.5)
---	-------

Um procedimento de otimização alternativo iterativo é necessário para se resolver o método *fuzzy c-means*. A execução deste procedimento requer a especificação do número de classes c e os pesos m . O algoritmo *fuzzy c-means* tem os seguintes passos:

1. Escolher $c(2 \leq c \leq n)$, $m(1 < m < \infty)$ e ε (um número pequeno para parar o procedimento iterativo) Mudar o contador $l=0$ e inicializar a matriz de pertinência $U(l)$.
2. Calcular o centro de classe, $v_i(l)$, usando a fórmula 3.4
3. Atualizar a matriz de pertinência $U(l+1)$ usando a fórmula 3.5 se $x_k \neq v_i(l)$. Caso contrário, mudar $\mu_{jk}=1$ se $j=1$ (ou $\mu_{jk}=0$ se $j \neq 1$).
4. Computar $\Delta = \|U^{(l+1)} - U^{(l)}\|$. Se $\Delta > \varepsilon$, incrementar l e voltar para o passo 2. Se $\Delta \leq \varepsilon$, o processo deve ser interrompido.

Esse grupo de algoritmos funciona melhor com séries temporais de igual comprimento porque o conceito de centros de classe não fica claro quando a mesma classe contém séries temporais de comprimentos diferentes.

Medidas de distância/similaridade

Uma chave importante na classificação é a função usada para medir a similaridade entre dois dados sendo comparados. Esses dados podem apresentar formas diferentes, incluindo valores de comprimento igual ou diferente, vetores de pares de valores de características, matrizes de transição, entre outros. A seguir são descritas as fórmulas de funções para definição de medidas de distância/similaridade, aspecto fundamental para a estruturação de operações de classificação.

Sendo x_i e v_j vetores p -dimensionais, a distância Euclidiana é computada pela fórmula 3.6:

$$d_E = \sqrt{\sum_{k=1}^P (x_{ik} - v_{jk})^2} \quad (3.6)$$

A distância do quadrado médio da raiz (ou distância geométrica média) é dada pela equação 3.7:

$$d_{rms} = d_E / n \quad (3.7)$$

A distância de Mikowski é uma generalização da distância Euclidiana, e é definida pela equação 3.8:

$$d_M = \sqrt[q]{\sum_{k=1}^P (x_{ik} - v_{jk})^q} \quad (3.8)$$

Na fórmula 3.8, q é um inteiro positivo. Uma versão normalizada pode ser definida se os valores medidos forem normalizados por divisão pelo valor máximo da sequência.

Sendo x_i e v_j vetores p -dimensionais. O fator de correlação de Pearson entre x_i e v_j , cc , é definido pela fórmula 3.9:

$$cc = \frac{\sum_{k=1}^p (x_{ik} - \mu_{ik})(v_{jk} - \mu_{vjk})}{Sx_i S v_j} \quad (3.9)$$

Onde μ_{x_i} e S_{x_i} são, respectivamente, a media e a dispersão de x_j , computados pela fórmula 3.10:

$$\mu_{x_i} = \frac{1}{P} \sum_{k=1}^P x_{ik} \quad \text{e} \quad S_{x_i} = \left[\sum_{k=1}^P (x_{ik} - \mu_{x_i})^2 \right]^{0.5} \quad (3.10)$$

Duas distâncias de correlação cruzada usadas comumente no algoritmo *fuzzy c-means* são definidas pela fórmula 3.11:

$$d_{cc}^1 = \left(\frac{1-cc}{1+cc} \right)^\beta \quad \text{e} \quad d_{cc}^2 = 2(1-cc) \quad (3.11)$$

Considerando cada intervalo de tempo como uma função linear, a distância STS é a soma da raiz das diferenças das inclinações em dois períodos de tempo a serem comparados. Matematicamente, a distância STS entre duas séries temporais x_i e v_j é definida pela fórmula 3.12:

$$d_{sts} = \sqrt{\sum_{k=1}^P \left(\frac{v_{j(k+1)} - v_{jk}}{t_{(k+1)} - t_k} - \frac{x_{i(k+1)} - x_{ik}}{t_{(k+1)} - t_k} \right)^2} \quad (3.12)$$

Onde t_k é o ponto no tempo para os pontos nos dados dado x_{ik} e v_{jk} . Para remover o efeito de escala, a estandardização z dos intervalos é recomendada.

Distorção dinâmica do tempo (DTW, Dynamic time warping em inglês) é uma generalização dos algoritmos clássicos para comparar seqüências discretas com seqüências de valores contínuos. Dados duas séries temporais, $Q = q_1, q_2, \dots, q_i, \dots, q_n$ e $R = r_1, r_2, \dots, r_j, \dots, r_m$, o DTW alinha os dois intervalos para que a diferença entre eles seja minimizada. Para estes casos, uma matriz $n \times m$ onde o elemento (i, j) da matriz contém a distância $d(q_i, r_j)$ entre os pontos q_i e r_j , onde é usada a distância Euclidiana. Um caminho de distorção, $W = w_1, w_2, \dots, w_k, \dots, w_K$ onde $\max(m, n) \leq K \leq m+n-1$, é um conjunto de elementos da matriz que satisfaz três condições: condição de fronteira, continuidade e monotonicidade. A condição de fronteira requer que o caminho de distorção comece e termine em células opostas diagonalmente em relação aos cantos da matriz. Isso é, $w_1 = (1, 1)$ e $w_K = (m, n)$. A condição de continuidade restringe os passos autorizados para células adjacentes. A condição de monotonicidade força os pontos do caminho de distorção a serem espaçados de maneira igual no tempo. O caminho da distorção tem a menor distância entre duas séries de interesse. Matematicamente:

$$d_{DTW} = \min \frac{\sum_{k=1}^K w_k}{K} \quad (3.13)$$

A Programação dinâmica pode ser usada para efetivamente achar esse caminho, ao avaliar a seguinte recorrência, que define a distância cumulativa como a soma da distância do elemento atual e do mínimo das distâncias cumulativas dos elementos adjacentes:

$$d_{cum}(i, j) = d(q_i, r_j) + \min\{d_{cum}(i-1, j-1), d_{cum}(i-1, j), d_{cum}(i, j-1)\} \quad (3.14)$$

A função de distância baseada em probabilidade para dados com erros foi originalmente desenvolvida para estudos de classificação de padrões sazonais. Foi definida a similaridade/distância entre duas estações, A_i e A_j , como a probabilidade de aceitar ou rejeitar a hipótese nula $H_0 : A_i \sim A_j$. Assumindo A_i e A_j , cada amostra independente compreendida por T elementos tirados das distribuições Gaussianas com médias x_{it} e x_{jt} , e desvio padrão σ_{it} e σ_{jt} , respectivamente, a estatística $\sum_{t=1, r} (x_{it} - x_{jt})^2 / (\sigma_{it}^2 + \sigma_{jt}^2)$ segue a distribuição quadrada com $r-1$ graus de liberdade. Consequentemente:

$$d_{ij} = x_{T-1}^2 \left(\sum_{t=1}^T \frac{(x_{it} - x_{jt})^2}{\sigma_{it}^2 + \sigma_{jt}^2} \right) \quad (3.15)$$

A hipótese nula $A_i \sim A_j$ denota $\mu_{it} = \mu_{jt}$ para $t = 1, \dots, T$.

Sejam P_1 e P_2 matrizes de probabilidades de transição de duas correntes de Markov (MCs, *Markov Chains*) com s distribuições de probabilidade cada e p_{1ij} e p_{2ij} serem a transição de probabilidade $i \rightarrow j$ em P_1 e P_2 . A distância assimétrica Kullback-Lieber de duas distribuições de probabilidade é definida por:

$$d(p_{1i}, p_{2j}) = \sum_{t=1}^T p_{1ij} \log(p_{1ij} / p_{2ij}) \quad (3.16)$$

A versão simétrica da distância Kullback-Liebler de duas distribuições de probabilidade é:

$$d(p_{1i}, p_{2j}) = [d(p_{1i}, p_{2i}) + d(p_{2i}, p_{1i})] / 2 \quad (3.17)$$

Onde a distância média entre P_1 e P_2 é então $D(P_1, P_2) = \sum_{i=1, s} D(p_{1i}, p_{2i})/s$.

Sendo $fr(\lambda_s)$ e $gr(\lambda_s)$ duas matrizes de estima espectrais, para duas séries diferentes de vetores estacionários com p dimensões e um número T de pontos no tempo, onde $\lambda_s = 2\pi s/T$, $s=1,2,\dots,T$. A J-divergência e a informação de divergência simétrica de Chernoff são computadas como descrito na fórmula 3.17 e 3.18:

$$J(fr; gr) = \frac{1}{2} T^{-1} \sum_s (tr\{f_T g_T^{-1}\} + tr\{g_T f_T^{-1}\} - 2p) \quad (3.18)$$

$$JB_\alpha(fr; gr) = \frac{1}{2} T^{-1} \sum_s \log \left(\frac{|\alpha f_T + (1-\alpha)g_T|}{|g_T|} + \frac{\log |\alpha g_T + (1-\alpha)f_T|}{|f_T|} \right) \quad (3.19)$$

Onde $0 < \alpha < 1$ e p é o tamanho das matrizes espectrais. Ambas as divergências são medidas de quasi-distância porque elas não satisfazem a propriedade da desigualdade dos triângulos.

Há uma versão estacionaria localmente da J-divergência para medir a discrepância entre duas séries temporais não estacionarias.

O índice de dissimilaridade baseado na função de correlação cruzada entre duas séries temporais é definido como:

$$d_{i,j} = \sqrt{\frac{(1 - p_{i,j}^2(0))}{\sum_{\tau=1}^{\max} p_{i,j}^2(\tau)}} \quad (3.20)$$

Onde $p_{i,j}^2(\tau)$ a função que denota a relação cruzada entre duas séries temporais x_i e v_j com retardação τ , e $\max$ é a retardação máxima. A contraparte de similaridade do índice acima pode ser definida como:

$$s_{i,j} = \exp(-d_{i,j}) \quad (3.21)$$

3.5. Considerações

Considerando a pesquisa teórica – metodológica exposta neste capítulo, pode-se observar a diversidade e complexidade de um processo de classificação, bem como a relevância do aprofundamento no assunto. Para proceder a opção de um algoritmo adequado e que contribua de forma decisiva na obtenção de conclusões e positivos para a solução do problema de classificação das regiões segmentadas das imagens de vídeo. Nos capítulos posteriores será tratada da importância da fusão dos métodos estatísticos, e de tratamento de imagem com os métodos de classificação para a separação dos objetos semânticos de maneira adequada.

Capítulo 4

Morfologia Matemática

4.1. Introdução

A morfologia matemática (ou morfologia) pode ser definida como uma teoria para a análise da forma e da estrutura. A morfologia é matemática no sentido em que é baseada em teoria dos conjuntos, geometria integral e álgebra de reticulados. Mas a morfologia matemática não é simplesmente uma teoria, mas é uma poderosa ferramenta para a análise e processamento de imagens.

A origem da morfologia matemática vem de meados dos anos 60 e a partir do estudo de materiais porosos de G. Matheron. Em seu trabalho pioneiro (*Eléments pour une théorie des milieux poreux* – 1967) Matheron propôs as primeiras transformações morfológicas. Quase que simultaneamente, em 1972, o desenvolvimento e a comercialização de um “*hardware*” especialista por J. C. Klein e J. Serra permitiu o uso dessas transformações para a solução de problemas práticos de análise de imagem. A partir dessas primeiras definições e formalizações de operações morfológicas para a investigação da geometria de objetos em uma imagem binária foi formalizada toda uma teoria por G. Matheron e J. Serra [FACON96].

De uma maneira geral, podem ser considerados dois tipos básicos de imagens às quais se aplicam as operações morfológicas: imagens binárias e em tons de cinza. A morfologia em cores ainda está em pleno desenvolvimento e não faz parte do escopo deste trabalho. Na morfologia aplicada às imagens binárias, para as quais os únicos valores de pixel possíveis são 0 ou 1 (branco e preto), na vizinhança de cada pixel da imagem original é testada uma configuração de pontos pretos e brancos (o elemento estruturante), de acordo com uma proposição específica, por exemplo: o elemento cabe no objeto ou toca o objeto.

Na morfologia aplicada às imagens em tons de cinza, além do procedimento descrito para as operações binárias é importante o conhecimento do valor dos pixels envolvidos tanto na imagem original, quanto no elemento estruturante, para a execução das operações de mínimo (ou ínfimo) e máximo.

4.2. Operações básicas

Podemos considerar as operações morfológicas como sendo comparações com um “elemento estruturante” escolhido de acordo com um conhecimento prévio das estruturas geométricas a serem analisadas e cuja finalidade é retirar ou alterar a forma dos objetos que estão relacionados (maiores, menores ou iguais) de alguma maneira a estes “elementos estruturantes”.

É possível classificar as operações morfológicas em três grandes grupos: filtragem, segmentação ou medidas de imagens. Todas essas operações são construídas através de dois operadores fundamentais da morfologia: a erosão e a dilatação. Esses operadores duais podem ser considerados como as duas pedras fundamentais da morfologia, já que todos os outros operadores mais complexos podem ser expressos em termos de dilatações e erosões.

O elemento estruturante é um conjunto conhecido e definido (em termos de forma e tamanho) usado para ser comparado ou medido na imagem a ser processada a partir de uma transformação. O formato e o tamanho do elemento estruturante permitem o teste e a quantização da maneira em que este está ou não contido na imagem. O elemento estruturante é freqüentemente referenciado como elemento estrutural plano quando ele é bidimensional e é normalmente utilizado para operações com imagens bidimensionais. De maneira similar, ele é referenciado como não plano, volumétrico ou de escala de cinza, quando se refere a elementos estruturantes compostos por pequenas imagens em tons de cinza, aplicáveis a operações em imagens em tons de cinza.

Os operadores morfológicos fundamentais necessitam da especificação de uma origem para cada elemento estruturante. De uma forma geral, pode se afirmar que a mudança na posição da origem num mesmo elemento estruturante resulta em uma translação do resultado final.

O formato e o tamanho dos elementos estruturantes devem ser adaptados às propriedades geométricas dos objetos das imagens a serem processadas. Por exemplo,

quando o interesse é a extração de objetos lineares em uma cena, elementos estruturantes lineares são os mais adequados. Na maioria dos casos os elementos estruturantes devem ser os mais simples possíveis e são escolhidos em função das propriedades de convexidade, não convexidade, isotropia e anisotropia, conforme exemplos de elementos estruturantes na figura 4.1.

Apesar de ser possível estabelecer tipos muito diversos de elementos estruturantes, somente um número limitado destes é usado em aplicações práticas. Alguns deles são:

- Segmentos de linhas – usados, normalmente, para extrair estruturas alongadas de imagens. Dois parâmetros importantes para este tipo de elemento estruturante são o tamanho e a orientação;
- Aproximação digital de discos – devido a sua isotropia, discos são elementos estruturantes muito utilizados. Várias aproximações de discos são possíveis como a cruz elementar e o quadrado elementar;
- Par de pontos – a erosão pelo par de pontos permite caracterizar o estado de dispersão de uma estrutura;
- Elementos estruturantes compostos – compreendem dois elementos estruturantes disjuntos que possuem a mesma origem;

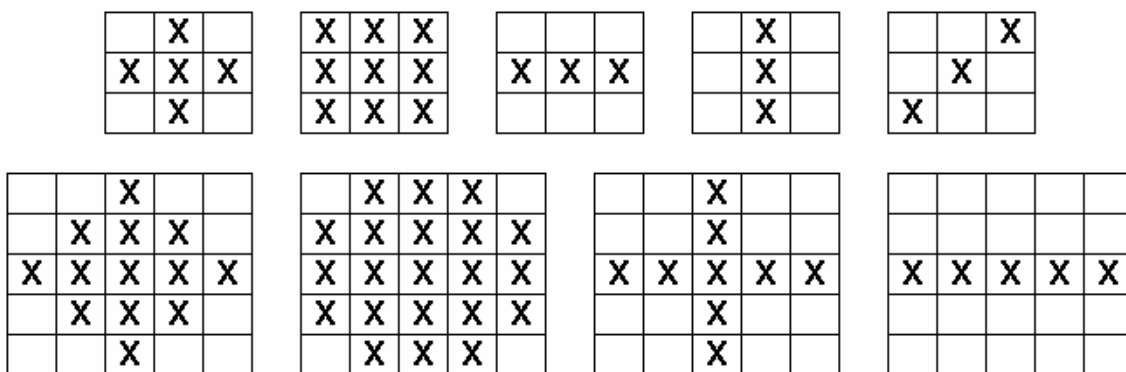


Figura 4.1 – Exemplo de elementos estruturantes 3x3 e 5x5.

4.3. Erosão

Uma das primeiras perguntas que podem ser feitas quando se compara um elemento estruturante e um conjunto ou objeto é: “O elemento estruturante está totalmente dentro do objeto?”. O objeto erodido é o lugar geométrico dos centros dos elementos estruturantes no qual esta resposta é afirmativa. Isto torna a erosão binária adequada para:

- Diminuir os objetos;
- Eliminar objetos menores que o elemento estruturante;
- Aumentar os buracos;
- Permitir a separação de objetos próximos;

Considerando X como o conjunto a ser erodido e B como o objeto estruturante escolhido, a erosão ($\mathcal{E}B$) pode ser definida como o lugar geométrico dos pontos x nos quais B é incluído em X , quando sua origem está posicionada em x :

$$\mathcal{E}_B(X) = \{x \mid B_x \subseteq X\} \quad (4.1)$$

A definição de erosão binária pode ser estendida diretamente em tons de cinza: a erosão de uma imagem f por um elemento estruturante B é definida como o mínimo ou ínfimo das translações de f pelos vetores $-b$ de B . Assim, o valor da erosão em um pixel x é o valor mínimo da imagem na janela definida pelo elemento estruturante quando sua origem está em x :

$$[\mathcal{E}_B(f)](x) = \min_{b \in B} f(x + b) \quad (4.2)$$

Isto torna a erosão em tons de cinza adequada para:

- Escurecer a imagem;
- Alargar e aprofundar os vales (padrões escuros);
- Conectar vales próximos;
- Reduzir/Eliminar picos; (padrões claros);
- Separar picos próximos;

Os valores de cinza de um elemento estruturante volumétrico devem ser escolhidos de conforme os valores de intensidade da imagem. Devido à dificuldade de se estabelecer essa condição antes do processo, e pelo fato de aumentarem as operações computacionais por conta da subtração, o elemento estruturante volumétrico é raramente utilizado na prática.

4.4. Dilatação

A dilatação é o operador dual da erosão e está baseada na seguinte pergunta: “O elemento estruturante toca o objeto em análise?”. Onde a resposta é afirmativa o objeto é acrescido do elemento estruturante, compondo assim o objeto dilatado. Isto torna a dilatação binária adequada para:

- Aumentar os objetos;
- Preencher pequenos buracos;
- Conectar objetos próximos;

Entende-se que a dilatação é o operador dual da erosão, pois, seus conjuntos têm entre eles uma correspondência um para um.

Considerando X como o conjunto a ser dilatado, B o elemento estruturante escolhido, a dilatação (∂) pode ser definida como o local dos pontos x onde B toca X quando sua origem coincide com x :

$$\partial_B(X) = \{x \mid Bx \cap X \neq \emptyset\} \quad (4.3)$$

Essa última definição pode ser estendida para imagens em tons de cinza: a dilatação de uma imagem f por um elemento estruturante B é definida como o máximo ou supremo das translações de f pelos vetores $-b$ de B :

$$[\partial_B(f)](x) = \max_{b \in B} f(x + b) \quad (4.4)$$

Isto torna a dilatação binária adequada para:

- Clarear a imagem;
- Alargar e engordar os picos (padrões claros);
- Conectar os picos próximos;

- Reduzir/Eliminar vales (padrões escuros);
- Separar vales próximos;

Diferente de imagens erodidas ou dilatadas por elementos estruturantes planos, as imagens dilatadas ou erodidas por elementos estruturantes volumétricos não estão limitadas pelos valores mínimos ou máximos dos tons de cinza da imagem de entrada. Em particular, a erosão com elemento estruturante em tons de cinza pode levar à imagens com resultados negativos, devendo ser usado com bastante cautela.

4.5. Propriedades da erosão e da dilatação

Com respeito à complementação, a dilatação e a erosão são transformações duais. Isso significa que qualquer erosão de uma imagem é equivalente à complementação da dilatação de uma imagem complementada com o mesmo elemento estruturante (e vice e versa). Em outras palavras, a erosão encolhe o objeto e expande o fundo, enquanto ocorre o oposto com a dilatação. Convém, entretanto, observar que a erosão e a dilatação não possuem transformação inversa, por não preservar a homotopia da imagem de entrada.

Em topologia, homotopia significa deformação de uma aplicação entre espaços topológicos [WIKIPEDIA06].

Espaços topológicos são estruturas que permitem a formalização de conceitos tais como convergência, conexidade e continuidade. Eles aparecem em praticamente todos os ramos da matemática moderna e são uma noção unificadora central. O ramo da matemática que estuda os espaços topológicos é chamado Topologia [WIKIPEDIA06].

Duas aplicações contínuas $f, g : X \rightarrow Y$ entre espaços topológicos são denominadas homotópicas se existir uma aplicação contínua $F : X \times [0,1] \rightarrow Y$, denominada homotopia, tal que $F_0 = f$ e $F_1 = g$, onde $F_t = f_{|X \times \{t\}}$.

A erosão e a dilatação são invariantes à translação. Essas operações também preservam a ordenação de objetos dentro das imagens, sendo assim, transformadas crescentes.

Com relação ao operador máximo, a dilatação é distributiva e a erosão é distributiva com relação ao operador mínimo.

A erosão de um conjunto com respeito a um elemento estruturante é sempre menor ou igual à dilatação realizada com o mesmo objeto estruturante [SOILLE99].

4.6. Operações geodésicas e reconstrução

Sejam x e y dois pontos quaisquer de um conjunto X . O arco geodésico de extremidade xy é definido como o arco de menor distância $dx(x,y)$ entre x e y , contido em X . Caso este arco não exista, esta distância é considerada infinita (figura 4.2). A função dx é denominada distância geodésica [FACON96].

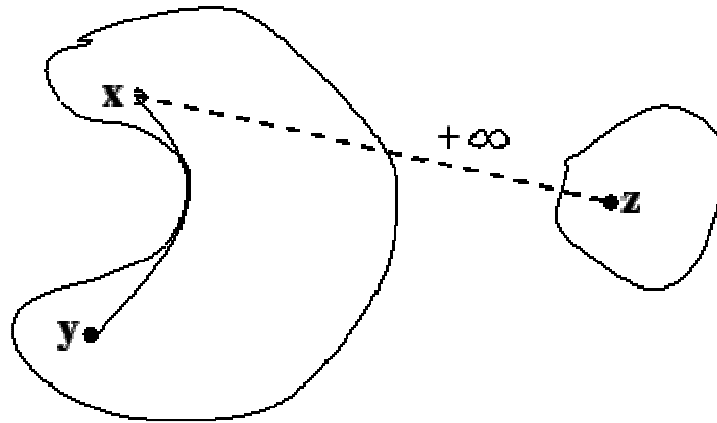


Figura 4.2 – Distância geodésica

As transformações morfológicas básicas podem ser definidas no espaço métrico (X, dx) . Por exemplo, se $Y \subset X$, pontos x de X tal que $B_x(x, \lambda)$ constituem o conjunto λ -dilatado de Y em X . Este conjunto é denotado por :

$$\delta_X^\lambda(Y) = \{x \in X : B_x(x, \lambda) \cap Y \neq \emptyset\} \quad (4.5)$$

onde $B_x(x, \lambda) = \{x \in X : dx(x, y) \leq \lambda\}$

Na prática, a dilatação geodésica de tamanho n , no espaço X , pode ser obtida a partir de n iterações da dilatação geodésica de tamanho 1 definida pela fórmula 4.6:

$$\delta_X^n(Y) = \underbrace{\delta_X^1(\delta_X^1(\dots \delta_X^1(Y)))}_{n \text{ vezes}} \quad (4.6)$$

A erosão geodésica de Y em X é dada por:

$$\mathcal{E}_X^\lambda(Y) = \{x \in X : B_x(x, \lambda) \subseteq Y\} \quad (4.7)$$

As transformações geodésicas são utilizadas na reconstrução de um conjunto conexo X demarcado por outro conjunto de referência Y, não vazio, denominado “marcador”. O conjunto X é a máscara binária que se pretende preservar. Uma dilatação geodésica de Y em X, até a idempotência, permite a reconstrução total desta componente. Como exemplo, consideremos o problema de eliminação de partículas parcialmente incluídas numa imagem (figura 4.3).

Supondo que $\partial \Sigma$ é o subconjunto de pontos que pertencem à fronteira de uma imagem Σ , então, o novo subconjunto X', que comporta as componentes incluídas totalmente em X, é dado por:

$$X' = X \setminus \delta_X^\infty(\partial \Sigma) \quad (4.8)$$

onde $\delta^\infty(.)$ indica a execução da dilatação até a idempotência.

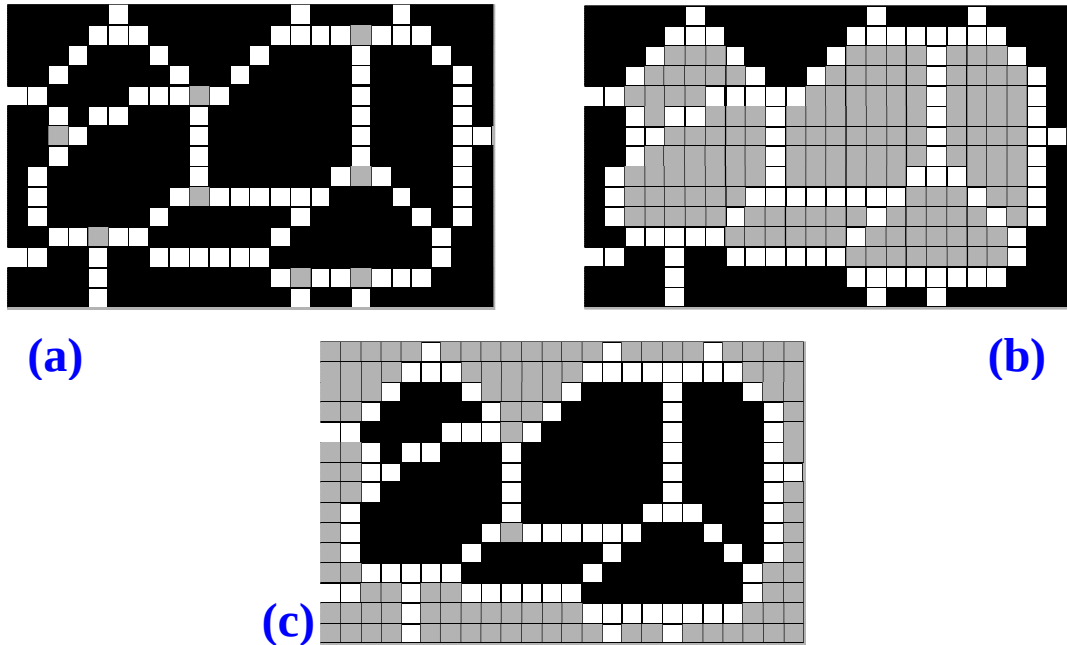


Figura 4.3 – Imagem Σ contendo partículas parcialmente incluídas (a); dilatação geodésica de $\partial \Sigma$ em X (b); subconjunto das partículas totalmente incluídas em Σ (c).

A noção de reconstrução geodésica é muito importante para o processamento morfológico, e constitui a base de inúmeras transformações mais complexas relacionadas, por exemplo, com separação de partículas, definição de operadores conexos e funções de extinção morfológicas.

No caso numérico, a dilatação geodésica de uma função y "marcando" uma função x ($y \leq x$) é dada por:

$$\begin{aligned} \delta_f^1(y) &= \min(\delta_1(y), x) \\ \delta_f^n(y) &= \underbrace{\delta_f^1(\delta_f^1(\dots \delta_f^1(y)))}_{n \text{ vezes}} \end{aligned} \quad (4.9)$$

A realização desta função até a idempotência define a reconstrução geodésica de x por y . A erosão de uma função $y \geq x$, $\varepsilon_x^n(y)$, é definida de maneira dual.

4.7. Considerações

Este capítulo apresentou a base para o uso de operações morfológicas nas transformações de imagens digitais binárias e em tons de cinza. Neste trabalho a morfologia foi amplamente usada, de modo que é essencial a sua compreensão. Com o uso das operações morfológicas será possível através de uma imagem binária contendo arquibancadas, placares, jogadores, e muito ruído, obter outra imagem binária com essencialmente apenas os jogadores e demais elementos presentes dentro de campo (como juiz, bandeirinhas, médicos, etc.).

Capítulo 5

Segmentação de imagens

Considerando o fato de que os pixels de uma imagem assumem diferentes significados conforme sua cor e posição relativa aos outros pixels, a segmentação de uma imagem é a separação destes pixels em grupos de significado distinto. Neste texto de estudo, considerar-se-á que os pixels de uma mesma região semântica se assemelham em termos de cor, e na sua distribuição espacial.

Normalmente a segmentação de imagens coloridas pode ser dividida nos seguintes grupos:

- Segmentação baseada no plano da imagem:
 - o Detecção de borda
 - o Reconstrução de região
- Segmentação baseada na análise do espaço de cor:
 - o Análise de histograma
 - o Análise de agrupamento

O processamento de vídeo demanda a adição do elemento “tempo”, que é uma informação auxiliar podendo ser usada em qualquer destes métodos. É entendida por tempo a seqüência dos quadros, que não necessariamente tem um tempo fixo entre eles, pois os vídeos podem ser codificados em várias taxas de quadros por segundo.

São encontrados ainda estudos de segmentação de vídeos comprimidos, ou melhor descrevendo, a segmentação é feita sem a necessidade da descompressão dos quadros individualmente. A vantagem é a velocidade, pois dispensa a descompressão, e podem ser aproveitadas estruturas do vídeo comprimido, como os vetores de

movimento, e os “*block averages*”, porém existem muitas limitações, como a dependência do tipo da compressão, e a qualidade do resultado final conforme os parâmetros da compressão [KOPRINSKA00].

Cada uma das abordagens citadas tem inúmeras derivações, com suas vantagens e suas limitações. O objetivo deste estudo é apresentar um novo tipo de segmentação, baseada em análise de histograma no espaço de cor HSV, independente de tempo, e sob imagens de vídeo já descomprimidas, mostrando seu potencial e suas possíveis aplicações.

Segmentação baseada no plano da imagem:

Detecção de borda: Detecta mudanças abruptas na propriedade de cor, [ZENZO86] propõe detectar os pixels de borda através de operadores diferenciais de primeira ordem. Já [CUMANI91] usa operadores diferenciais de segunda ordem, ambas as técnicas são consideravelmente sensíveis a ruídos, e trazem a dependência de uma imagem bem focada e com bordas bem definidas, o que não costuma ser realidade em imagens esportivas, em que são freqüentes os casos de planos mais distantes e fora de foco em uma mesma imagem.

Reconstrução de região: Normalmente baseado em processos de “*region growing*” [TRÉMEAU00], que é o agrupamento de pixels por semelhança de cor a partir de regiões iniciais, conhecidas como sementes, ou “*split and merge*” [PANJAWANI95], que considera também a textura, na qual as regiões são fundidas conforme sua semelhança, e separadas conforme sua discrepância, considerando os atributos espacial e de cor dos pixels. Em ambos os casos existem problemas para a definição de limiares de semelhança para o agrupamento.

Segmentação baseada na análise do espaço de cor:

Análise de histograma: Pode trabalhar com vários espaços de cor, [OHTA80] parte a imagem em regiões através da análise recursiva do histograma, [SHAFARENKO98] aplica o algoritmo *watershed* em um histograma cromático bidimensional para a construção de classes. Estes métodos são normalmente sensíveis a ajustes de parâmetros.

Análise de agrupamento: [TAJKAHASHI95] e [ISMAIL89] propõe o uso de técnicas de agrupamento, como *c-means* e *isodata*, que são métodos muito usados para a análise de agrupamento. Porém, o bom desempenho destes métodos exige uma boa

escolha do centro dos grupos. [SOUTO00] utiliza o crescimento de regiões para segmentação de imagens de matiz, porém ressalta que uma mesma imagem pode ter diferentes resultados conforme a localização das sementes. São usados vários parâmetros, como tamanho mínimo da região, da classe, e os limiares de reclassificação e de segmentação, características próprias do tipo de imagem a ser processada.

[RODRIGUES03] em seu trabalho sobre recuperação de informação visual, compara o uso de: cor, forma e textura juntas e separadamente. O autor conclui que o uso das três informações ao mesmo tempo quando as três apresentam um resultado bom é melhor do que o melhor resultado das três, porém quando qualquer uma delas tem um resultado ruim, o seu uso conjunto é pior do que o melhor resultado individual. Isto traz dúvida sobre a aplicação dos atributos em conjunto para a segmentação das imagens.

Portanto, a maior parte das técnicas demonstradas na literatura depende de heurísticas e parâmetros, o que inviabiliza grande parte das aplicações.

Capítulo 6

Imagem e Vídeo

6.1. Qualidade

Mensurar uma distância, uma massa ou uma velocidade normalmente não é uma tarefa difícil, pois há uma referência, um instrumento de medição calibrado, ou seja, é um procedimento elementar, porém, mensurar qualidade de vídeo é uma tarefa muito diferente. A forma de mensurar qualidade de vídeo tem mudado recentemente com a adoção cada vez maior do vídeo digital.

Existem duas formas de se avaliar a qualidade de uma sequência de vídeo, a objetiva e a subjetiva.

A forma objetiva consiste em modelos matemáticos que emulam os resultados subjetivos da avaliação de qualidade, baseados em critérios e métricas que podem ser medidos objetivamente. Os métodos objetivos são classificados conforme a disponibilidade do sinal de vídeo original, que supostamente está em alta qualidade. Desta maneira, podem então ser classificados em:

- Métodos de referência completa
- Métodos de referência parcial
- Métodos sem referência

A maneira mais clássica de avaliar a qualidade do vídeo digital é a contagem da relação PSNR (*peak signal-to-noise ratio*)(Formula 6.1), que nada mais é que a relação sinal/ruído entre o vídeo original e o vídeo já processado. Esta técnica é normalmente usada em CODECs de vídeo populares, como o DivX, XviD, etc. A PSNR é então uma

métrica objetiva de qualidade, não assistida, e que pode ser implementada computacionalmente, conforme descrito na fórmula 6.1 . Atualmente métricas mais complexas e mais precisas estão sendo desenvolvidas. A SSIM (*Structural SIMilarity*) e a VQM (*Video Quality Metric*) são exemplos de novos métodos [CSMSU06].

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} ||I(i, j) - K(i, j)||^2$$

$$PSNR = 10 \cdot \log_{10} \left(\frac{MAX_I^2}{MSE} \right) = 20 \cdot \log_{10} \left(\frac{MAX_I}{\sqrt{MSE}} \right) \quad (6.1)$$

A maneira subjetiva é a mais tradicional para avaliar a qualidade de uma imagem, e consiste em uma avaliação estimada, feita pelo melhor sistema de visão e percepção existente, o ser humano. Porém este sofisticado sistema de visão e percepção é altamente subjetivo, não podem ser encontrados dois idênticos, e são altamente influenciados por inúmeros fatores, como cansaço, humor, uma doença degenerativa da visão ou do cérebro. Isto incorre em uma margem de erro muito grande, logo, a solução adotada para minimizar este erro é utilizar vários visualizadores, e fazer uma média de suas opiniões (fórmula 6.2). Esta técnica é conhecida por MOS (*Mean Opinion Score*), ou Contagem Média de Opinião. Outro problema sério deste tipo de abordagem é o fato de necessitar de visualizadores treinados, e mesmo com visualizadores treinados o resultado final pode variar consideravelmente [WANG02]. Os passos para uma avaliação subjetiva são os seguintes:

- Escolher a sequência de vídeo para teste;
- Definir os ajustes do sistema que se pretende avaliar;
- Escolher o método de teste (como as sequências são apresentadas aos especialistas como as opiniões são coletadas);
- Convidar um número suficiente de especialistas (recomenda-se pelo menos 15);
- Realizar o teste;
- Calcular a média das opiniões dos especialistas para cada sequência de vídeo;

6.2. Compressão

A compressão de dados, como sugere o próprio nome, é o ato de comprimir dados, ou seja, torná-los menores através de diversos algoritmos de compressão, reduzindo a quantidade de *bits* para representar uma determinada informação, sendo ela uma imagem, um texto, ou um arquivo qualquer.

Um trabalho que pode ser citado para ilustrar as diferenças entre os formatos de compressão é o “*MSU Subjective Comparison of Modern Video CODECs*” [CSMSU06], onde 50 especialistas do CS MSU Graphics & Media LAB Vídeo Group da Rússia avaliam subjetivamente quatro populares CODECs através do método SAMVIQ, que é um método subjetivo de análise de qualidade [KOZAMERNIK04].

Os CODECs utilizados para este teste são descritos na tabela 6.1.

CODEC	Fabricante	Versão
DivX	DivX Networks	6.0 b1571-CenterOfTheSun
XviD	Open Source project	1.1.-125 (“xvid-1.1.0-beta2”)
X264	Open Source project	Core 48 svn-352M by Sharktooth
WMV	Microsoft Corporation	9.0

Tabela 6.1 – CODECs utilizados

Todos os CODECs foram ajustados para obter *bitrates* de 690 kbps e 1024 kbps, sendo que, quatro cenas de vídeo de aproximadamente 10 segundos foram comprimidas. As duas primeiras em 704x288 pixels e as duas segundas em 720x416 pixels.

Os testes são limitados, pois cada um dos 50 examinadores deve ver cada seqüência em cada *bitrate* pelo menos uma vez.

$$MOS_k = \sum_{i=1}^{\text{experts_num}} \frac{\text{mark}_{i,k}}{\text{experts_num}} \quad (6.2)$$

Onde “k” é o número da seqüência para o qual o MOS é calculado, $\text{mark}_{i,k}$ é a nota dada pelo especialista “i” para a cena “k” da seqüência, e experts_num é o número de especialistas.

O resultado da comparação subjetiva pode ser conferido na figura 6.1 e na tabela 6.2.

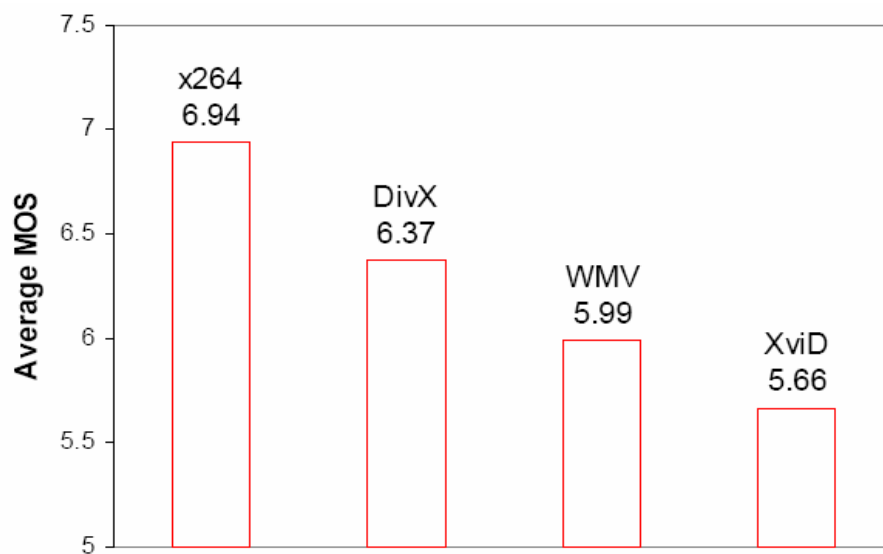


Figura 6.1 – Resultado da comparação MOS

Codec / <i>bitrate</i>	Colocação
x264 1024	1
DivX1024	2
XviD 1024	3
x264 690	4
WMV1024	5
DivX 690	6
WMV 690	7
XviD 690	8

Tabela 6.2 – Colocação na avaliação de qualidade dos especialistas, da melhor para a pior qualidade.

Foram feitas também avaliações objetivas, nos métodos PSNR, VQM e SSIM, e o resultado destes métodos foram comparados com as avaliações subjetivas, trazendo um grau de correlação conforme a tabela 6.3. Como pode ser visto, dentre as métricas avaliadas, o SSIM é a alternativa mais próxima de uma avaliação subjetiva [CSMSU06].

Métrica	Correlação com o MOS
PSNR	0,802
VQM	0,729
SSIM	0,937

Tabela 6.3 – Correlação entre MOS e outras técnicas objetivas, onde 1 reflete 100% de similaridade.

Na conclusão dos trabalhos, pode ser observado que o x264 é o algoritmo que melhor preserva os detalhes da imagem, sendo superior até mesmo às soluções proprietárias.

Outra observação importante é que nem sempre as medidas objetivas refletem os mesmos resultados que as subjetivas, e o mais impressionante, as avaliações subjetivas são mais precisas que as objetivas (quando elaboradas com rigor científico-metodológico, como é o caso do trabalho acima mencionado).

Neste trabalho, todos estes fatores e conceitos foram considerados para criar uma métrica subjetiva-objetiva, capaz de mensurar a qualidade não dos vídeos, mas do resultado de seu processamento, e principalmente que pudesse ser factível com os recursos de tempo e pessoal disponíveis.

6.3. CODECs de vídeo

Um codec de vídeo é um dispositivo ou módulo de software que permite o uso da compressão de dados para o vídeo digital. Normalmente são empregados métodos de compressão com perda de informação. Em vídeo digital existe uma relação complexa entre qualidade de vídeo e a quantidade de informação necessária para representá-lo (conhecida como *bitrate*).

Existem dois tipos de CODECs:

- Sem perdas (*lossless*, em inglês)
- Com perdas (*lossy*, em inglês)

CODECs sem Perdas:

Os CODECs sem perdas são aqueles que codificam som ou imagem, com o objetivo de comprimir o arquivo sem desvirtuar o conteúdo original. Ao ser

descomprimido, o novo arquivo será idêntico ao original. CODECs deste tipo normalmente geram arquivos codificados que são entre 2 a 3 vezes menores que os arquivos originais. Emissoras de televisão e rádio utilizam-no com frequência para manter a qualidade do som ou imagem.

Como exemplos desse tipo de codec encontram-se os seguintes: shorten, flac, wavpack e monkey's audio, para som. Para imagem: lossless MJPEG, HuffYUV e FFmpeg Video 1.

CODECs com Perdas:

Os CODECs com perdas são CODECs que codificam som ou imagem, com a finalidade de alcançar maiores taxas de compressão, gerando certa perda de qualidade. A perda de qualidade, então, é equilibrada com a taxa de compressão para que não venham a ser criados artefatos percebíveis.

Quanto ao som, por exemplo: se um instrumento muito baixo toca ao mesmo tempo em que outro instrumento mais alto, o primeiro poderá ser suprimido na compressão, uma vez que dificilmente será ouvido. Quando da ocorrência deste tipo de caso, somente um ouvido bem treinado pode identificar a supressão do instrumento.

Em relação aos vídeos, a informação suprimida também é a menos relevante, como, detalhes, e a alta frequência. Cita-se como exemplo um vídeo de telejornal, quando a parte do fundo, que é mais estática, tende a perder os detalhes de pequenos movimentos, privilegiando o jornalista que se mexe com mais frequência.

Taxa de *Bits*:

A taxa de *bits* (do inglês, *bitrate*) é uma das medidas da qualidade de um arquivo comprimido com um codec com perdas. A taxa de *bits* representa o tamanho final desejado para o arquivo e, normalmente, é apresentada como *kbits/s*.

1 *kbit/s* significa que a cada segundo, o codec tem 1000 *bits* do arquivo final para utilizar, ou seja, se um arquivo de som tem 8 segundos e é comprimido a uma taxa de 1 *kbit/s*, o arquivo final terá 8 *kbits* ou 1 *kbyte*. Logo, quanto maior a taxa de *bits*, melhor será a qualidade do arquivo final, uma vez que o codec terá mais espaço para poder comprimir o arquivo original, necessitando de descartar menos "detalhes" do arquivo.

São exemplos de CODECs de vídeo:

H.261: Usado em produtos de videoconferências e videotelefonia, foi o primeiro padrão prático de compressão digital, e foi desenvolvido pela ITU-T.

MPEG-1 Part 2: Usado em VCDs (*video* CDs) e em algumas aplicações de vídeo *online*. Tem qualidade próxima a de um VHS.

MPEG-2 Part 2: Usado em DVD, SVCD, e na maioria dos sistemas de transmissão de vídeo digital via ar e cabo.

H.263: Usado em produtos de videoconferências, videotelefonia e internet. Foi um passo significativo na compressão de vídeo *progressive scan* (não entrelaçado).

MPEG-4 Part 2: Padrão MPEG para ser usado na internet, transmissão, e armazenamento;

MPEG-4 Part 10: É o estado da arte do padrão de compressão MPEG, e está sendo adotado rapidamente para uma variedade de aplicações;

DivX, XviD, 3ixv: Implementações diferentes do MPEG-4 Part 2;

WMV: Família de CODECs da Microsoft, usada desde vídeos em baixa resolução para internet até HDTV;

Realvideo: Desenvolvido pela RealNetworks, foi um codec popular, mas está perdendo importância rapidamente;

Capítulo 7

Estado da Arte

Nos últimos anos, a evolução dos meios de armazenamento e de geração de conteúdo digital, provocou demanda por soluções eficientes para lidar com este conteúdo estimulou a pesquisa e desenvolvimento de várias abordagens distintas no processamento de imagens. [DUARTE02] em seu trabalho de recuperação de imagens digitais por similaridade cromática, demonstra a eficiência do espaço de cor HSV quando o objeto de atenção é a cor. No espaço de cor HSV a percepção cromática é mais uniforme, e possibilita uma análise independente de fatores como luminosidade irregular, degradês, sombras, etc. Porém, pela simplicidade da abordagem, o seu uso não traz resultados animadores em aplicações reais, sendo necessária a junção de outras técnicas mais elaboradas.

Já [RODRIGUES03] em seu trabalho também sobre recuperação de imagens visuais, usa três abordagens diferentes, a primeira utilizando informações do relacionamento espacial entre as regiões da imagem, além de histograma de cores, mapa de bordas e também informações de textura. A segunda, usando decomposição em valores singulares para extrair informações de co-ocorrência de características, transformando o espaço original em um novo espaço, chamado de “espaço latente”, e por fim usa um modelo bayesiano. Este modelo bayesiano combina o resultado de ordenação de métodos clássicos, obtendo um desempenho ainda melhor. De seu trabalho, cabe ressaltar a idéia de que em um processo de reconhecimento, maior quantidade de informações não significa maior certeza, pois enquanto algumas informações podem ser muito oportunas, outras podem trazer confusão, influenciando mais negativamente do que positivamente no resultado final. Ele cita como exemplo um sistema de classificações de orquídeas, onde instintivamente pode se pensar em usar a

informação de cor ou textura, estas informações na verdade trarão colapsos no processo de detecção, que de maneira bem mais adequada pode usar a forma como informação básica para classificação. Outra contribuição deste autor é tentar aproximar os sistemas de visão computacional dos princípios da visão biológica humana, que de fato é o melhor exemplo disponível. Para isso ele procura complementar as informações de cor, forma e textura, com informações “contextuais”, baseado no agrupamento de informações primitivas por semelhança, e por contexto. Importante manifestação do autor também é a constatação da precariedade de trabalhos que dão a devida atenção a estes itens.

Em seu trabalho de classificação de imagens aéreas, [SOUTO00] compara um método baseado em matiz, contra outro baseado em isoseg (que é um classificador não supervisionado), mostrando a superioridade do uso da matiz em especial quando a quantidade de informação é menor. Informação neste caso significa “bandas espectrais de imagem aérea”. O desempenho usando matiz com mais bandas espectrais se deve especialmente por similaridade cromática dos elementos a serem diferenciados (como por exemplo, pasto e eucaliptos). Seu trabalho traz ainda aspectos importantes a serem observados quando se usa a matiz para segmentação de regiões de uma imagem: um deles é o risco de ter uma classificação errada quando existe similaridade cromática, nestes casos é preciso avaliar a possibilidade do uso de outros elementos como forma ou textura; o outro aspecto fundamental é o comportamento instável do atributo matiz sob valores muito pequenos de saturação e/ou intensidade. Por problemas especialmente de captação e conversão, em regiões tendendo ao preto, o atributo matiz pode tomar valores discrepantes, não afetando no aspecto visual, mas no processamento que considera a matiz como única referência.

Em tempo, o fato do padrão para armazenamento e transmissão de vídeo digital ser o RGB exige que haja um processo de conversão de espaço de cor em cada quadro do vídeo. Este processo é um pouco custoso computacionalmente.

[WANG01] em seu trabalho de segmentação de imagens, traz várias informações importantes, pois trata especialmente de imagens com baixa profundidade de campo. A profundidade de campo é a extensão da área da imagem que aparece totalmente no foco. Ela está relacionada à distância de focalização (longe ou perto), ao tipo de objetiva (distância focal) e à abertura do diafragma. Quanto maior a distância em relação ao objeto, maior será a profundidade de campo. Quanto maior a abertura do diafragma, menor será a profundidade de campo, ou seja, menor será a área da imagem

que estará totalmente no foco. Para resumir, é a distância da câmera que está aceitavelmente bem focada na imagem. Se por um lado uma imagem com o fundo desfocado pode ser perfeita para uma segmentação baseada em textura, por outro traz sérios problemas na segmentação utilizando a matiz. Em uma partida de futebol, muitas vezes nem toda a região visível está com foco adequado, por este motivo a cor do campo se funde com a cor dos jogadores. Mesmo que visualmente seja possível entender que o jogador esteja na posição determinada, o processo de segmentação possivelmente não se fará com a mesma precisão.

Recentemente, novas abordagens para a segmentação de vídeo utilizaram a análise de apenas uma imagem que representa todo o segmento de vídeo. Esta imagem de representação do vídeo é chamada de ritmo visual [GUIMARÃES01]. Estas técnicas diferenciam-se das demais por não utilizar medidas baseadas na dissimilaridade de quadros. Elas são baseadas na busca por padrões em uma imagem criada a partir de uma amostragem de cada quadro de uma sequência de vídeo, preservando muitas das características do vídeo original. A amostragem de cada quadro é obtida através da extração de uma fatia diagonal, vertical ou horizontal do quadro. Uma imagem é criada utilizando estas amostras e, desta forma, o conteúdo do vídeo sofre uma simplificação, pois cada fatia é transformada numa faixa vertical da imagem. Vale constar que este método é eficiente para segmentações mais simples, onde são detectadas mudanças representativas no fluxo das imagens, e também é rápido computacionalmente. Existe, porém, nessa abordagem, algumas questões que demandam uma melhor reflexão, como o problema da transição de imagens. Transições muito bruscas normalmente confundem os métodos de segmentação que usam informações dos quadros anteriores, e transições lentas, onde uma imagem vai se transformando gradualmente em outra, costumam complicar a maioria dos métodos que não são feitos para detectar especificamente este tipo de transição.

[SANTOS04] trata de maneira mais completa a variação das cenas do vídeo, através de intersecção de histogramas dos quadros. Esta técnica pode inclusive detectar a variação gradual de imagens, pois o histórico do histograma mostra a transição de maneira gradual também, todavia, esta variação não pode ser vista de maneira pontual, mas sim, efetuando uma análise da diferença entre um histograma antes e depois desta variação, para então obter a mesma medida de variação tida em um corte brusco. Esta técnica é conhecida por *twin-comparision*. [SANTOS04] desta forma contribui

mostrando de maneira eficaz o potencial do uso do histograma como ferramenta para a análise estatística de imagens.

[ANDRADE03] trata o rastreamento de objetos baseado no domínio de regiões, usando informações relacionais ao invés de absolutas. E trata as inconsistências entre quadros com uma busca numa árvore de figuras. Para localizar os objetos são usados vetores de movimento, que se baseiam na diferença entre quadros para identificar a movimentação de blocos de pixels. Esta abordagem é computacionalmente mais intensiva, e, conforme o autor, ainda imatura para uso prático. Além do que, a técnica exige também o conhecimento prévio de alguns fatores, não podendo ser considerada completamente automática.

Para o problema das transições graduais em vídeo, [DREW00] apresentou uma interessante abordagem. É baseada apenas em operações elementares, pois usa basicamente intersecção de histogramas, que é um método computacionalmente muito eficiente e pouco custoso. Uma equação que tem como entrada os dados dos histogramas das imagens quadro a quadro apresenta valor constante caso esteja havendo uma transição. O problema é que quando a cena inicial (antes da transição) e a final (após a transição) apresentam muito movimento, esta técnica fica comprometida, pois não apresenta resultados satisfatórios.

Encontra-se ainda no trabalho de [WANG02] a utilização de informações globais da imagem, e homogeneidade local, baseando-se sempre no componente matiz no espaço de cor HSV. Aspectos relevantes nesse método devem ser observados, como: o fato de não haver uma técnica de segmentação absoluta, que serve em qualquer imagem em qualquer condição, e a maior simplicidade da segmentação por cores no espaço HSV (em especial no caso do uso do algoritmo de deslocamento médio), mesmo com a característica cíclica do componente H. Uma característica importante é ser um método não supervisionado, e que não precisa de conhecimento prévio do número de regiões. O algoritmo de deslocamento médio já havia sido utilizado com sucesso em outras abordagens de agrupamento, segmentação de imagens, rastreamento, etc., porém, em todos os casos citados apenas utilizavam informações globais, negligenciando as informações locais. Logo [WANG02] propõe a introdução da medida de homogeneidade local no algoritmo de deslocamento médio. Os resultados são consideráveis se comparando com outros trabalhos nos quais é feita utilização apenas de informação global.

[CABROL05] propõe uma técnica de segmentação de imagens por cor em tempo real, usando classificação de cores e agrupamento de regiões. As aplicações deste tipo de técnica são várias, o autor, no entanto, concentra-se no uso para competições de robôs. O processamento em tempo real quer dizer na verdade que todos os quadros adquiridos pelo conjunto de câmera/digitalizadora são processados. Os passos para a classificação das regiões começam com um agrupamento hierárquico piramidal, um agrupamento das regiões piramidais, um agrupamento por cor, baseado em uma classificação de regiões, que pode ser subdividida em classificação de cor das regiões, e um agrupamento baseado em classificação de cores. Os resultados em termos de tempo e desempenho são melhores que os da maioria dos métodos comparados. O grande benefício, no entanto, que torna o algoritmo tão rápido é o pouco acesso a pixels, graças ao agrupamento hierárquico, e a classificação de cor por regiões baseada no valor médio.

[GU04] em seu trabalho sobre segmentação de tomadas de vídeo de futebol americano, usa um modelo adaptativo para detecção da cor do campo. Para definir que uma imagem mostra o campo, são procuradas linhas brancas (que são características do campo de futebol americano) cruzando a imagem. Neste ponto a abordagem já se torna muito limitada para uso em outros jogos, realizados em campos muito diferentes em suas marcações. O modelo adaptativo para a separação do campo requer um ajuste inicial que é a definição do parâmetro da cor verde, o qual vai sendo ajustado conforme os quadros vão sendo processados para compensar diferenças de iluminação, cor da grama, sombras etc. Parâmetros também são utilizados para definirem quão grandes e contínuas devem ser as linhas do campo. Para a detecção das linhas é usado o algoritmo de Canny (para fazer uma detecção de bordas), seguido de uma transformada de Hough (que localiza primitivas na imagem, neste caso as linhas). Obviamente o resultado deste processo depende muito da boa definição do que é o gramado e de quão bem este gramado é separado do resto dos outros elementos. Apesar do modelo adaptativo usado por [GU04] ser suficiente para a aplicação proposta, ele é visivelmente insuficiente para segmentar de maneira mais detalhada os jogadores e demais elementos da imagem. Isto fica claro na imagem ilustrativa da figura 8.6b, onde aparecem o campo em preto e os outros elementos em branco, pois comparando o quadro original com o limiarizado, é possível perceber que muitos jogadores foram confundidos com campo, e várias partes do campo não foram marcadas como tal.

[VANDENBROUCKE03] utiliza uma abordagem original para melhorar o método de classificação de pixels na segmentação. É escolhido um novo espaço de cor através de componentes pertencentes aos diversos espaços de cores clássicos, que tem características muito diferentes entre si. Logo a idéia principal é classificar pixels em um espaço de cor especialmente desenvolvido para ter melhor rendimento na discriminação das classes de pixels. Este espaço de cor é chamado “espaço de cor híbrido adaptado”, e é criado a partir de uma seqüência de características selecionadas de maneira supervisionada. Os testes foram feitos baseados em imagens de jogos de futebol, e as cores das camisas foram usadas para determinar os times. Os resultados são consideráveis, e demonstram a forte influência do espaço de cor na análise da imagem. Por outro lado, o fato de precisar converter cada imagem em vários espaços de cor para extrair as componentes necessárias, e a necessidade de supervisão, torna o método pouco robusto, em especial para múltiplas aplicações, e para processos automatizados, além de ser muito custoso computacionalmente.

O autor menciona que nenhum espaço de cor clássico é bem adaptado para discriminar todas as classes de jogadores. Menciona ainda que o espaço de cor híbrido torna o processo de classificação mais rápido que nos outros espaços de cor, porém, esta afirmação sem dúvida desconsidera as partes iniciais do processo, que são a conversão da imagem do seu espaço de cor original, e a seleção das componentes mais adequadas para o espaço de cor híbrido. Entretanto, o método se mostrou eficaz na detecção de jogadores que na imagem aparecem parcialmente sobrepostos, o que é uma deficiência conhecida de todas as aplicações de segmentação.

[NEEDHAM01] trabalha exatamente no problema de rastreamento de jogadores mesmo sob oclusão, usando como exemplo o rastreamento de jogadores de futebol em quadras internas. Para segmentar o fundo, o autor utiliza subtração de *background*, e isto requer uma câmera fixa (além do auxílio do ambiente controlado de uma quadra interna), por isso, a técnica se torna muito restrita para aplicações mais complexas, que utilizam câmeras em movimento, diferentes escalas, e iluminação não controlada. A base do trabalho de [NEEDHAM01] é a condensação, que é uma técnica que permite a propagação condicional de densidades ao longo do tempo, e tem sido usada com sucesso em rastreamento de objetos mesmo sob oclusão. O espaço de cor HSV é usado para segmentar o fundo e os jogadores, mas o motivo não é citado claramente. A identificação dos jogadores é feita por “forma”, ou seja, modelos são criados e depois localizados no vídeo. O resultado final foi de 28% a 56% das trajetórias identificadas,

porém o resultado na identificação dos jogadores visualmente foi fraco, pois os exemplos mostrados no artigo nunca tinham mais que a metade dos jogadores identificados.

[EKIN02] escreve sobre detecção da região de cor dominante, com aplicação em análise de vídeos esportivos. A proposta do trabalho é a detecção mesmo em condições de variação das condições de luminosidade ao longo do tempo. A robustez do algoritmo se deve a adaptação da estatística da cor dominante no espaço de cor primário para o auxílio usando um espaço de cor secundário e a fusão destes dois espaços. Os resultados em ambos os espaços de cor são fundidos para uma melhor precisão, eliminando ruídos. Os resultados demonstrados através das imagens mesmo na fusão dos dois espaços de cor não trazem muita precisão na identificação da área dominante. Em especial o espaço de controle traz muita discrepância na identificação do campo. Os resultados práticos demonstraram também que o sistema sofre muito com a interferência de fatores externos mais fortes, como variações bruscas de luminosidade.

[MEHROTRA03] em seu trabalho sobre sumarização e análise automática de vídeos de futebol usa uma abordagem baseada no espaço de cor HSI. Suas contribuições incluem além de detecção da cor dominante, algumas funções de mais alto nível, como detecção do juiz, de gols, segmentos em “*slow-motion*”, e detecção da grande área. Para as funções de mais alto nível a análise semântica usa cinemática e características baseadas em objeto (mesmo que computacionalmente mais pesado).

A detecção da cor dominante é aprendida no início do vídeo e então é automaticamente atualizada para se adaptar às variações temporais. A cor dominante é definida como o valor médio de cada componente de cor que é computado em volta do respectivo pico de histograma. São usados picos de histograma da componente H de várias imagens sucessivas, e os pontos de corte à direita e à esquerda do pico são definidos como o ponto onde a contagem de pixels está em 20% do valor obtido no pico. O problema deste tipo de consideração é que nem sempre ele representa a realidade, em especial quando existem na imagem outras regiões com cores parecidas, ou quando a região do campo não é tão grande em relação a outras. As imagens que mostram a segmentação da região dominante demonstram muita discrepância e falsas detecções de objetos ao longo do campo, possivelmente, resultado de algumas suposições e valores fixos sobre a análise do histograma. A análise adaptativa traz informações de quadros anteriores, o que pode ser prejudicial em especial após

transições um pouco mais bruscas, isto também pode explicar a deficiência da detecção do campo.

[FIGUEROA04] utiliza para o rastreamento de jogadores uma abordagem por representação por grafos. Para tal abordagem são utilizadas múltiplas câmeras, tratando a oclusão com a divisão das regiões segmentadas quando possível, com o auxílio da cor da camisa. Os nós representam os jogadores, e as arestas as distâncias entre eles. O algoritmo de segmentação usado segue os seguintes passos: Extração de *background* com auxílio estatístico para atualização da região de fundo; filtros morfológicos para eliminar ruídos, limiarização da imagem, e rotulação dos pixels conectados. A técnica apresentada requer para o rastreamento dos jogadores em todo o campo um mínimo de 4 câmeras.

[THIRDE03] ao tratar a questão da segmentação de objetos semanticamente relevantes, usa uma abordagem espaço temporal probabilística, com modelagem gaussiana de informação cromática, e um esquema de segmentação bayesiana combinando informação espacial e cromática. Os resultados são adequados para várias aplicações, em especial a compressão de vídeo, todavia, são necessárias intervenções do usuário, definindo os quadros chave, e três parâmetros para iniciar o processo.

[OZYILDIZ02] usa uma abordagem interessante, fundindo a informação de cor e de textura para segmentar e rastrear objetos em uma seqüência de vídeo. Apresenta baixo custo computacional, permitindo o rastreamento em tempo real. O problema encontrado nesta abordagem é a sensibilidade às mudanças de *background*, e a mudança de textura de objetos, que ao se moverem rapidamente, tem sua textura alterada por arrasto, problema este intrínseco aos equipamentos de captura de imagens. Em imagens esportivas, onde o fundo varia constantemente, e os objetos se movem rapidamente, este tipo de abordagem não propicia resultados animadores.

[NAKAMURA99] usa redes neurais e segmentação por cor para rastreamento de objetos. O modelo utilizado é o *fuzzy ART*, que é uma rede neural para aprendizado competitivo. O espaço de cor escolhido é o YUV, e outro espaço de cor criado para facilitar o tratamento pela rede neural. A vantagem no uso de *fuzzy ART* é o aprendizado não supervisionado, e a possibilidade de uso em um contexto não estacionário. Este modelo funciona bem para imagens simples, como cenários de futebol de robôs, porém em vídeos mais complexos a técnica não se mostra satisfatória.

[LEFÈVRE02] usa uma abordagem de segmentação multi-resolução por cor para a extração de *background* de imagens externas. A metodologia baseia-se na criação

de uma representação piramidal da imagem original, que é então re-amostrada e segmentada em várias resoluções usando a informação cromática. O espaço de cor HSV foi escolhido para a tarefa, pois após a comparação dos vários espaços de cor foi o que mostrou maior robustez em especial quanto às mudanças de iluminação. O autor usa o exemplo da segmentação de vídeos de futebol para mostrar a eficiência do algoritmo proposto. Uma das suposições do autor para a aplicação do método é a uniformidade da cor de fundo na imagem em baixas resoluções, o que nem sempre é realidade, em especial em ambientes de iluminação artificial irregular, sombra causada por nuvens, ou projetadas pela própria arquitetura do estádio. A conclusão do próprio autor é que a técnica pode ser usada em um pré-processamento para definir a posição inicial dos objetos no cenário.

Capítulo 8

Metodologia Proposta

Do ponto de vista temporal, existem basicamente duas abordagens para processamento de imagens de vídeo, uma delas usa a informação dos “n” quadros anteriores, e a outra se limita a trabalhar com o quadro atual, independente de informações anteriores. Cada técnica tem suas vantagens e desvantagens, e nada impede que sejam usadas em conjunto.

Análises de fluxo ótico, rastreamento de objetos, detecção de oclusão, segmentação de tomadas de vídeo, etc., analisam as informações ao longo do tempo, necessitando para o início do funcionamento um número determinado de quadros para sua inicialização. Este número varia de uns poucos quadros, até vários minutos de vídeo, conforme a técnica.

Para ser independente de informações de quadros anteriores, cada quadro deve ser processado em sua totalidade, não impedindo que otimizações sejam feitas na extração das informações. Por exemplo, em seqüências de vídeo muito parecidas, é possível o aproveitamento do histograma de quadros anteriores, porém isto não é alvo deste estudo.

A técnica aqui proposta busca ser independente do tempo, não requerendo informações anteriores para processar um quadro de vídeo. Isto possibilita o processamento não só de vídeo como de imagens estáticas. Este tipo de abordagem permite que os quadros sejam processados desde o primeiro até o último, e a análise não é penalizada por variações bruscas na imagem, como seria, por exemplo, no caso de subtração de *background*. A abordagem proposta também pode ser usada para auxiliar outras técnicas dependentes do tempo, fornecendo as informações semânticas necessárias para interpretar a imagem estática e o vídeo em mais alto nível.

A base deste estudo é uma nova técnica de detecção automática da região dominante, que tem por objetivo ser usada em imagens esportivas, para o processamento de mais baixo nível, que é separar o fundo da imagem da parte semanticamente relevante. Um dos objetivos desta técnica é dispensar interações do usuário ou quaisquer tipos de parâmetros, de forma a poder ser utilizada em aplicações não supervisionadas, como o processamento automático embarcado.

O algoritmo principal recebe como entrada uma imagem esportiva, no caso deste estudo, jogos de futebol, e tem como saída uma imagem binária, com a região dos jogadores em branco e o fundo em preto, bem como a definição das regiões, separando os jogadores dos diferentes times.

A metodologia consiste nos seguintes passos:

1. Converter a imagem para o espaço de cor HSV;
2. Fazer um estudo estatístico do canal H
 - a. Extrair seu histograma;
 - b. Localizar o maior pico deste histograma;
 - c. Detectar no histograma as regiões ao redor do pico que ainda contém informações relacionadas com ele;
 - d. Localizar os limiares de cor do campo;
3. Limiarizar a imagem H, com 2 limiares correspondendo às cores máxima e mínima do campo;
 - a. É definido que o campo recebe a cor preta;
 - b. Os jogadores e outros elementos do campo recebem a cor branca;
4. Eliminar as arquibancadas e demais elementos indesejados de fora do campo através de operações matemáticas e morfológicas;
5. As regiões restantes sobre o fundo preto são rotuladas e classificadas;

8.1. Conversão RGB -> HSV

A grande vantagem do uso do espaço de cores HSV é a independência do canal da H (Hue) em relação aos demais. Para otimizar o processo, a conversão do espaço de cor pode ser feita apenas parcialmente, pois apenas o elemento matiz (H) do HSV é necessário para o restante do processo.

Grande parte das implementações de funções de conversão e bibliotecas de processamento de imagens, comprimem o canal H em 8 *bits*, quando o necessário seriam 9 *bits*. O componente H comumente é representado por cores que vão de 0° a 360°, no entanto, com 8 *bits* se fazem apenas 256 combinações, enquanto com 9 *bits* são possíveis 512 combinações para comportar as cores. Esta compressão está relacionada com o padrão adotado pelos computadores, que define a representação de um *byte* por 8 *bits*, e com o fato de toda a arquitetura computacional estar desenvolvida para tratar informação desta maneira. Logo, para manipulação de uma componente de cor (e não para seu armazenamento), é necessário um múltiplo de um *byte*, e as opções mais lógicas seriam 16*bits*, ou 8 *bits*, numa delas desperdiçando 7 *bits* e complicando as operações matemáticas, e na outra perdendo informação. A escolha feita normalmente é comprimir a informação em 8 *bits*, (em especial quando não se necessita de uma precisão muito grande, e a operação inversa não será feita), representando as cores do H de 0° a 360°, por inteiros que vão de 0 a 180.

8.2. Estatística sobre o canal H

A inovação desta técnica está na forma automática de definir os limiares de cor da região de interesse. É neste ponto que o algoritmo decide quais pixels pertencem e quais não pertencem à determinada região. O algoritmo se baseia em identificar a largura exata do pico. Isto se dá da forma como segue:

Definição de variáveis:

- Seja $H = \{H(i, j)\}$, que denota o canal matiz H de uma imagem digital colorida, onde $h = H(i, j)$ indica a matiz no pixel (i, j) da imagem; O canal matiz H de uma imagem digital colorida é formada de 360 valores inteiros, então, qualquer h pertence ao conjunto $\{0, 1, \dots, 360\}$;
- P_H representa o pico do histograma do canal matiz H;

- P_{esq} representa a base esquerda do pico do histograma do canal matiz H;
- P_{dir} representa a base direita do pico do histograma do canal matiz H;

O algoritmo foi desenvolvido de maneira a identificar os pontos P_{esq} e P_{dir} , de modo que a região que vai do ponto P_{esq} até P_{dir} deve conter a porção do histograma que representa a região de interesse, ou seja, o a região do pico do histograma que contém a informação de campo.

Início do algoritmo:

1. Identifica-se o pico do histograma P_H do canal matiz H da imagem;
2. O histograma é dividido em 2 sub histogramas H1 e H2, sendo que:
 - a. O sub-histograma H₁: tem seu início no primeiro “180 éximo” ponto antes do pico de histograma P_H (que é exatamente no meio do histograma);
 - b. O sub-histograma H₂: tem seu início em P_H e seu fim no último “180 éximo” ponto depois do pico de histograma P_H ;
3. São criados dois (2) sub-histogramas acumulados $S_1(h)$ e $S_2(h)$, a seguir:

$$S_1(h) = \sum_{j=h}^{P_H} P_1(j) \qquad S_2(h) = \sum_{j=PH}^{P_H+h} P_2(j) \qquad (8.1)$$

O somatório das probabilidades $P_1(j)$ do sub-histograma H₁, para $j = h, h+1, \dots, P_H$, onde $0 \leq h \leq P_H$.

O somatório das probabilidades $P_2(j)$ do sub-histograma H₂, para $j = P_H, P_H+1, \dots, P_H+h$, onde $P_H \leq h \leq 360 - P_H$.

4. Calcula as derivadas primeiras de $S_1(h)$ e $S_2(h)$, gerando duas variáveis $D_1(h)$ e $D_2(h)$, seguindo a fórmula $D_n(h) = S_n(h+1) - S_n(h)$, onde $n = 1, 2$.
5. Localiza os maiores picos das derivadas primeiras $D_1(h)$ e $D_2(h)$, respectivamente M_1 e M_2 ;
6. Localiza a partir de M_1 e M_2 as passagens por zero (*zero-crossing*) mais próximas de cada pico, respectivamente chamadas M_{01} e M_{02} ;
7. Localiza a partir de M_{01} e M_{02} os limites do pico, respectivamente P_{esq} e P_{dir} ;

Fim do algoritmo

Em todos os pontos do algoritmo a característica cíclica do componente H é respeitada, de modo que não haja discrepâncias nos valores de limiar definidos.

Definidos M_{o1} e M_{o2} é então identificado no histograma exatamente o pico [P_{esq} e P_{dir}] ou ainda a sua largura ($P_{dir} - P_{esq}$) (figura 8.2).

A figura 8.1 (a) apresenta uma imagem de uma partida de futebol, de resolução 640x480, digitalizada a partir de uma transmissão de TV, comprimida com divX 5.2.1. A tabela 8.1 mostra os valores máximo/mínimo do canal H para pixels do campo (onde H tem valores de 0-360°).

A primeira passagem por zero depois do pico da derivada D_1 está no ponto $M_{o1} = 8$, e a primeira passagem por zero depois do pico da derivada D_2 está no ponto $M_{o2} = 14$, como pode ser observado nas figuras 8.3 e 8.4. Logo, é entendido que a região do pico do histograma H, que compreende os pixels do campo, se estende desde a cor 144° até a cor 166°, como pode ser observado nas figuras 8.3 e 8.4. Estes valores são obtidos através das fórmulas $PH - M_{o1}$ e $PH + M_{o2}$, sendo que $PH = 152$.

Os valores resultantes são semelhantes aos obtidos manualmente (através de um software editor de imagens), e que constam na Tabela 8.1. Para adquirir os valores manualmente foi utilizada a ferramenta “conta gotas” do software editor de imagem, e foi se deslocando o cursor pela imagem, especificamente na área do campo, a procura dos maiores e menores valores de H. A diferença entre os valores obtidos manualmente e os obtidos automaticamente, se deve à pouca quantidade de pixels das cores mais extremas da região de pico do histograma, que são descartadas por não serem representativas. Para ilustrar o resultado, na figura 8.1b está a mesma imagem, porém com os pixels de componente H de 144° até 166° demarcados em preto.

Detecção	H Min	H Máx
Automática	144	166
Manual	138	171

Tabela 8.1 – Detecção de campo manual/automática

Exemplo:



Figura 8.1 – (a) Imagem de partida de futebol, (b) Imagem com cores detectadas como gramado em preto

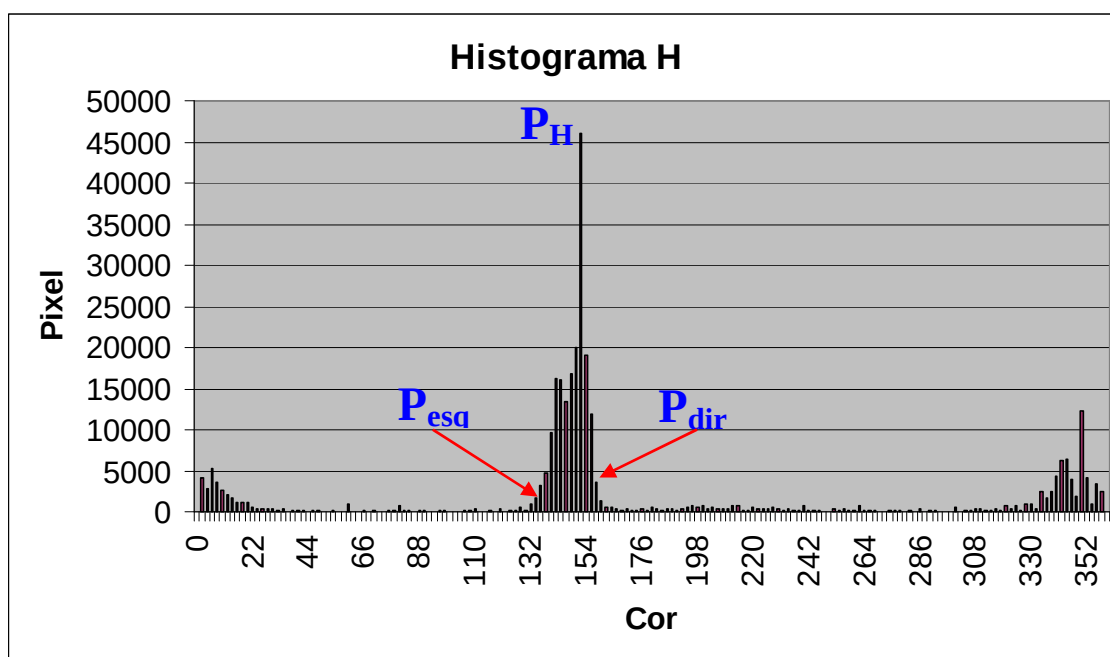


Figura 8.2 – Histograma do canal H da Figura 8.1a

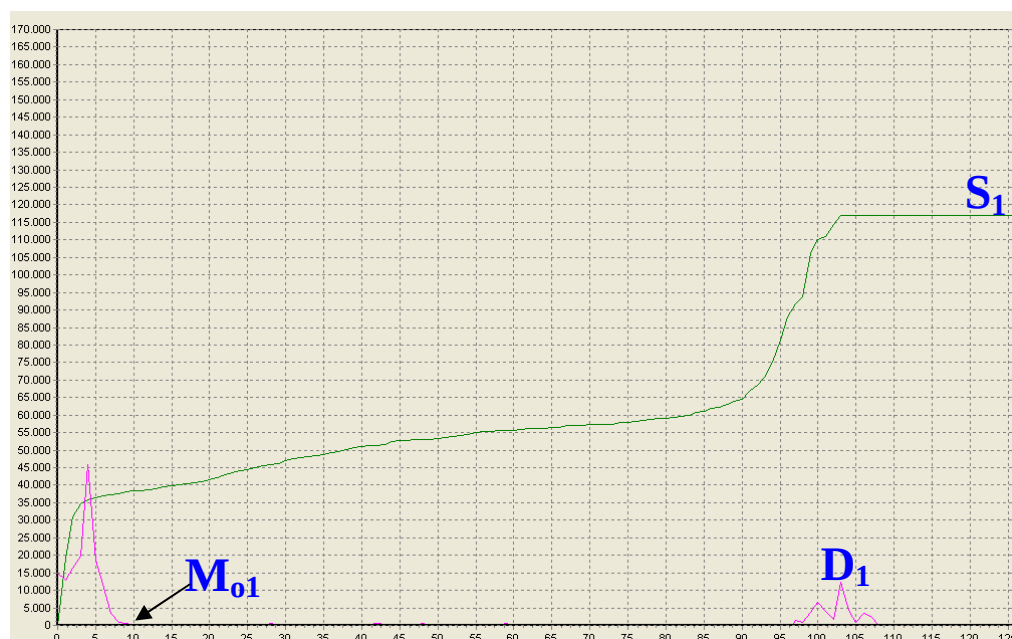


Figura 8.3 – (S_1) acumulação de histograma, (D_1) derivada de S_1 , (M_{o1}) primeiro ponto onde D_1 se torna zero depois do pico

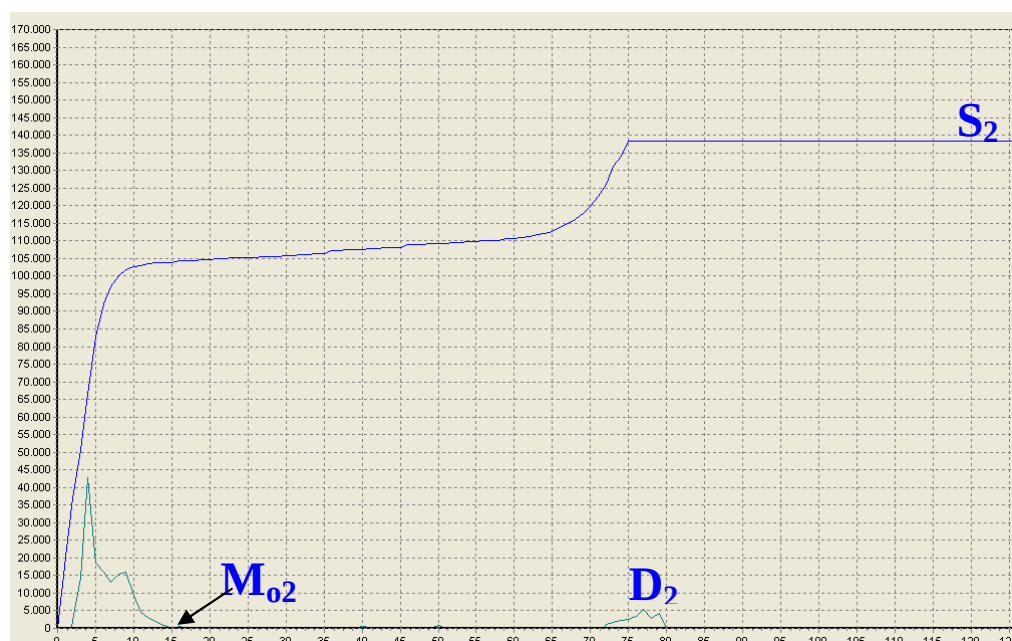


Figura 8.4 – (S_2) acumulação de histograma, (D_2) derivada de S_2 , (M_{o2}) primeiro ponto onde D_2 se torna zero depois do pico

8.3. Limiarização (binarização)

Uma vez com as informações sobre a faixa de cores pertencente ao campo, é preciso, de alguma maneira, separar o campo dos objetos semanticamente relevantes. A limiarização não resolve este problema por completo, porém, é um passo fundamental

para a qualidade do resultado final, uma vez que seu resultado é a eliminação da região de cor dominante na imagem, restando os vários outros objetos da cena, como as arquibancadas, placas de publicidade, etc.

A binarização com dois limiares tem por objetivo representar a faixa entre os limiares com uma cor, e as demais tonalidades com outra. Neste caso, a faixa de cor entre os limiares será representada pelo preto, e as demais pelo branco. Esta operação é feita sobre uma imagem de 8 *bits*, representando o canal H da imagem original.

Mais uma vez é preciso cuidado para considerar a característica cíclica do canal H, pois os limiares podem não estar necessariamente em ordem do ponto de vista horizontal. Supondo dois limiares, “L1” e “L2”, definidos respectivamente nos ângulos 350° e 30° , logo a faixa de cor selecionada não será da $30^\circ - 350^\circ$, mas sim de $350^\circ - 360^\circ$ e mais $0^\circ - 30^\circ$. Nos campos de futebol onde a cor predominante é a verde, não ocorrem problemas desta natureza, mas como pode ser visto na figura 8.5, campos com predominância da cor vermelha podem ter faixas de limiar não contínuas, ou seja, um campo contido em uma faixa de cor de 40° , poderia estar distribuído nas faixas de 340° a 360° e 0° a 20° , por exemplo.

Exemplos de uma imagem limiarizada podem ser vistos na figura 8.6, onde a imagem (a) é normal, e a imagem (b) é a limiarizada. As regiões em preto estão dentro da faixa de cores identificada como campo, e o resto das cores está em branco.

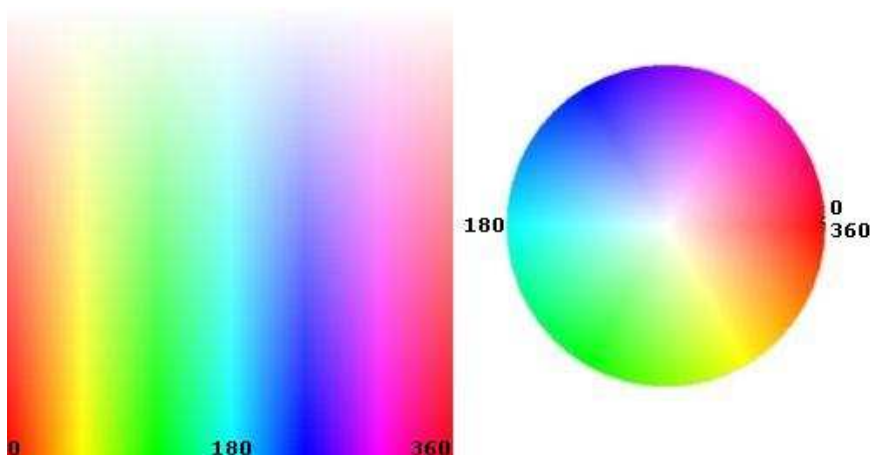


Figura 8.5 – Espaço de cor HSV de forma horizontal (a) e de forma circular (b)

8.4. Morfologia Matemática e Ferramentas Matemáticas

As operações matemáticas e morfológicas sobre uma imagem são processos demorados, pois operam sobre todos os pixels da imagem por repetidas vezes. Do ponto

de vista computacional, são várias leituras, processamentos, e gravações na memória por cada *byte*. Por outro lado, as operações efetuadas são elementares, muitas vezes feitas em blocos de pixels, ocupando apenas um ciclo de *clock* por bloco, graças às instruções multimídia que existem nos processadores mais novos.

O processo de eliminação das informações não desejadas da imagem parte do princípio que os jogadores estão em branco, isolados pelo campo em preto. As outras informações em branco nas extremidades da imagem se tratam de: arquibancada, letreiros de publicidade, placar, ou outras imagens inseridas na transmissão.

Para ser flexível, trabalhando com uma variedade de tamanhos, qualidades, e tipos de vídeo, foi necessário desenvolver uma rotina baseada no menor número de premissas possível, evitando ser muito destrutiva, preservando ao máximo os jogadores, e eliminando todo o resto. Toda premissa restritiva pode melhorar a qualidade do resultado final, porém restringe sua aplicabilidade. Um dos objetivos principais deste projeto é ser não supervisionado, e não paramétrico, pois de outra forma não seria uma boa base para outras pesquisas, e nem tampouco traria uma contribuição científica sólida dentro de sua proposta.

Nas operações morfológicas descritas, sempre que omitido, o elemento estruturante é o quadrado 3x3, com uma iteração.

O processo adotado consiste nos seguintes passos:

1. É criada uma imagem preta de mesmo tamanho que a imagem binarizada;
 - Conserva as características de dimensão e profundidade de cor;
2. É feita uma moldura em branco da espessura de um pixel na imagem criada;
 - Consiste em pintar de branco as primeiras e últimas linhas e colunas da imagem;
3. É efetuada a operação de reconstrução binária sobre a imagem de moldura, tendo a imagem do canal H filtrada e binarizada como máscara;
 - A operação de reconstrução vai preservar na imagem apenas os elementos fundidos com a borda, eliminando todo o resto;
4. A imagem reconstruída no passo 3 é então dilatada;
 - O processo de dilatação tem por objetivo fundir partes ilhadas na região de arquibancada, (possivelmente por conter tonalidades iguais à do campo), e deixar rastros na imagem final;

5. É criada uma nova imagem;
 - Novamente com as mesmas dimensões e profundidade de cor que as outras;
6. A nova imagem recebe a imagem binarizada subtraída da imagem dilatada, e deve conter apenas os jogadores em branco em um fundo preto;
 - É um processo de subtração absoluto, sem valores negativos, os valores negativos são truncados para zero.
7. Esta nova imagem é erodida com o elemento estruturante quadrado;
8. Depois é reconstruída, tendo a imagem do passo número 6 como máscara;
9. Depois de reconstruída a imagem é então dilatada com o elemento estruturante “linha vertical” para fundir pedaços de jogador eventualmente separados;

A figura 8.6 mostra uma seqüência de imagens com o resultado destas operações, onde:

- a. Imagem original;
- b. Imagem limiarizada onde as regiões identificadas como campo estão em preto, e as demais em branco;
- c. Imagem reconstruída a partir de uma imagem completamente preta com bordas de espessura de 1 pixel da cor branca;
- d. Imagem “c” dilatada
- e. Imagem “b” subtraída da imagem “d”
- f. Imagem “e” erodida;
- g. Imagem “f” reconstruída com a mascara “e”
- h. Imagem “g” dilatada

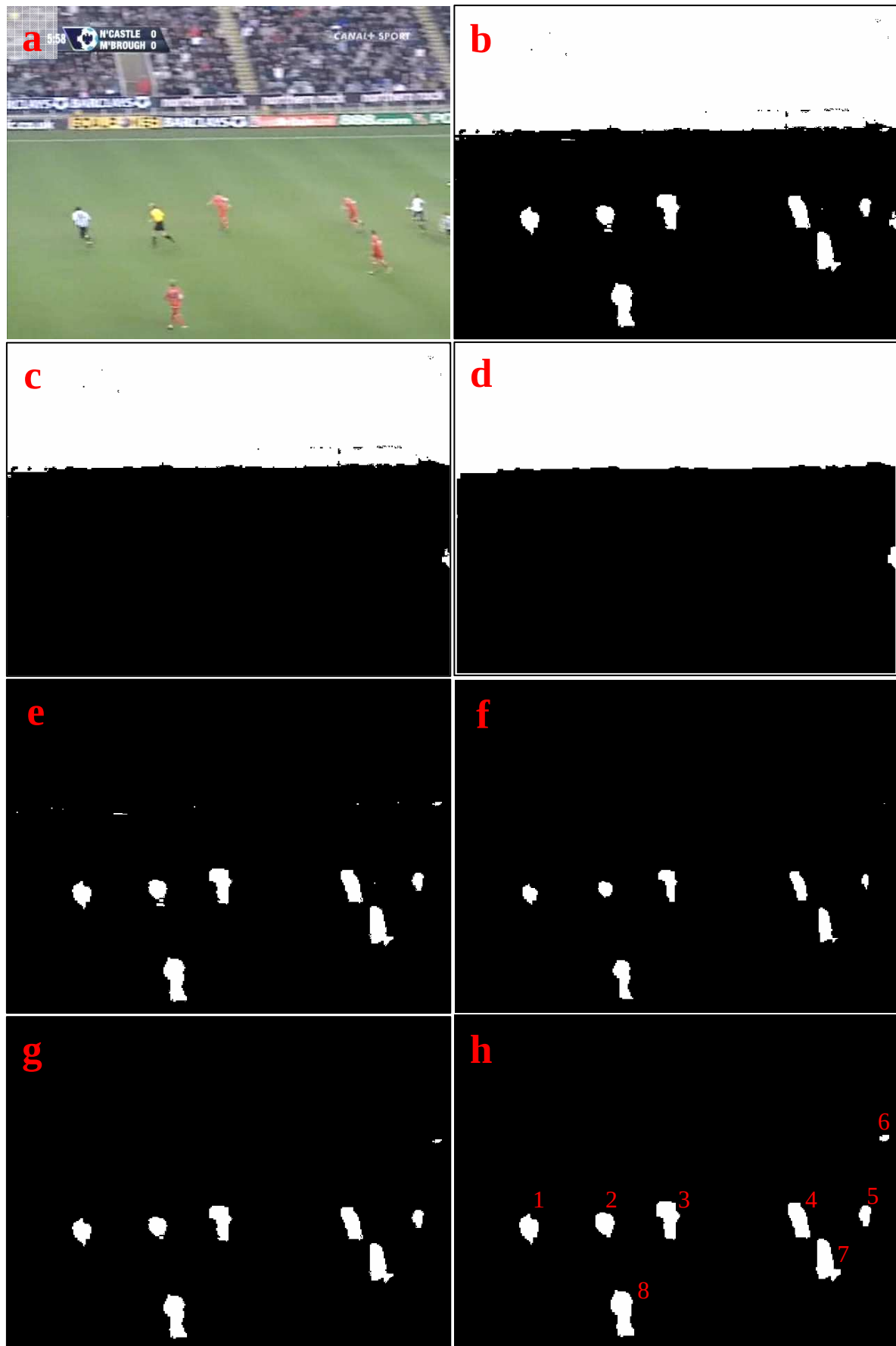


Figura 8.6 – Sequência de operações morfológicas

8.5. Agrupamento

A separação dos grupos apesar de parecer uma tarefa elementar traz várias complicações a serem consideradas. A primeira delas é que como o trabalho se propõe a não ser supervisionado, a classificação dos grupos deve ser automática, isto implica em identificar dois fatores:

- Quantos grupos existem na imagem?
- Qual região pertence a qual grupo?

Em primeira instância são dois grupos principais, representados pelos jogadores de ambos os times, porém, tem-se bem mais do que isto:

- Jogadores do Time 1
- Jogadores do Time 2
- Juiz
- Goleiro do Time 1
- Goleiro do Time 2
- Regiões de campo não identificadas como tal

O juiz normalmente usa cor diferente dos jogadores, bem como os goleiros, que não devem parecer com nenhum dos outros jogadores e árbitros, sendo necessários grupos para cada um deles.

Em algumas partes da imagem, regiões de campo não são identificadas como campo, este problema é causado por muitos motivos, tanto físicos quanto por características do próprio algoritmo. Alguns dos fatores que podem supostamente causar este problema são:

- Campos de má qualidade (com áreas sem grama);
- Barro;
- Poças de água;
- Regiões de tonalidades consideravelmente diferentes do resto do campo;
- Etc.

Logo, o número máximo de grupos possíveis na imagem está entre zero e “n” para o caso do futebol, onde “n” é dependente da complexidade da imagem, e do número de regiões segmentadas. Cada quadro de um vídeo pode ter um número diferente de regiões e de grupos. Em suma, a tarefa número um é, tendo as regiões segmentadas, identificar o número de grupos formados por estas regiões.

A tarefa número dois é classificar as regiões, identificando-as como pertencentes a cada um dos grupos.

Basicamente, o processo de agrupamento envolve a identificação de um conjunto de grupos que descrevam o conjunto de dados. Neste sentido, o processo de agrupamento, que é uma técnica de aprendizado não supervisionado, tem por objetivo maximizar a homogeneidade entre os objetos de um mesmo grupo e, ao mesmo tempo, maximizar a heterogeneidade entre objetos de grupos distintos. O conceito de similaridade (ou homogeneidade) normalmente envolve o conceito de distâncias entre pares de objetos.

Considera-se que o processo de agrupamento envolve a separação de um conjunto X de objetos em uma coleção de objetos mutuamente disjuntos C_i de X. Formalmente, considere-se um conjunto de “n” objetos $X = \{x_1, x_2, \dots, x_n\}$ que deve ser agrupado, onde cada $x_i \in \mathcal{R}^\rho$ é um vetor de atributos composto por “ ρ ” medições reais (contínuas). Os objetos devem ser agrupados em grupos não sobrepostos $C = \{C_1, C_2, \dots, C_k\}$ onde k é o número de grupos, tal que:

$$C_1 \cup C_2 \cup \dots \cup C_k = X, \quad C_i \neq \emptyset, \quad e \quad C_i \cap C_j = \emptyset \quad \text{para } i \neq j \quad (8.2)$$

O número de possibilidades de se classificar “n” objetos em “k” grupos é dado por:

$$N(n, k) = \frac{1}{K!} \sum_{i=0}^k (-1)^i \binom{k}{i} (k-i)^n \quad (8.3)$$

Supondo que se tem em uma determinada imagem 25 objetos, para classificá-los em 5 grupos formam-se: $N(25, 5) = 2436684974110751$ combinações. Desta maneira fica fácil perceber a complexidade de agrupar corretamente os conjuntos de objetos de um jogo.

A metodologia empregada neste trabalho procura otimizar do número de grupos, definindo um valor de k que maximize a homogeneidade interna dos grupos e a heterogeneidade entre os grupos distintos. Para isto, é fundamental a definição da função de distância a ser utilizada entre os elementos.

A tendência natural é separar as regiões pela sua cor, e uma vez que o espaço de cor HSV está sendo usado, a matiz (H) é a forma mais simples e eficaz de se trabalhar com a cor.

No decorrer da pesquisa, várias tentativas foram feitas no sentido de ter uma referência de cor para cada região segmentada, mas quase todas apresentaram um péssimo resultado, entre elas:

- A média das cores de cada região (usando o canal H);
- Mediana das cores de cada região (usando também o canal H);
- Ambas as técnicas acima, porém com filtros de pixels com cor próxima à do campo;
- Combinações destas técnicas, porém usando os canais RGB;

O problema encontrado por qualquer uma destas técnicas é a pouca quantidade de informação existente em cada região, em alguns casos menos de uma dezena de pixels por região. Somado a isto, existe ainda o problema da influência das cores adjacentes à região (exemplificado na figura 8.7), compressão, etc. Isto causa um fenômeno pouco percebido, que é o agrupamento de pixels de cores intermediárias nas regiões segmentadas. Isso dificulta todo e qualquer cálculo considerando todas as cores da região, pois este fenômeno torna as regiões que são muito diferentes, mais parecidas entre si, tanto pela coloração tendendo à cor do campo, como pela perda da informação de cor realmente relevante, neste caso, da cor da camisa dos jogadores.

A solução então deve desconsiderar apenas parte da informação da região segmentada, contabilizando apenas o que é útil para mensurar a proximidade das regiões em termos de cor, possibilitando um agrupamento sem discrepâncias.

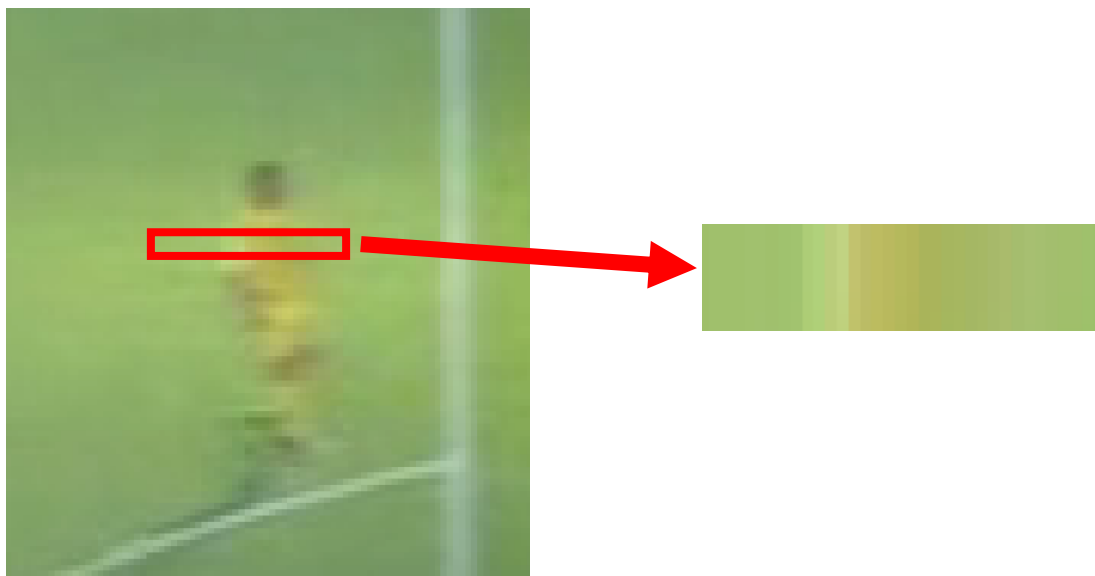


Figura 8.7 – Influência das regiões adjacentes na imagem: (a) Imagem Original, (b) Linha ampliada.

A técnica desenvolvida neste trabalho para determinar o número de regiões em cada cena e agrupar as regiões é descrita a seguir:

Definição de variáveis:

- Seja “n” o número de regiões segmentadas;
- R_n representa uma região segmentada pertencente ao conjunto de “n” regiões segmentadas;
- G_n representa o grupo de cores mais significativo da região R_n
- HR_n representa o histograma da região R_n
- $V_C(n)$ representa o vetor de valores centrais da região mais representativa do histograma HR_n
- OV_C representa o vetor VC ordenado do menor para o maior elemento;
- D_{OVC} representa o vetor contendo a derivada primeira do vetor OVC ;
- mD representa a média do vetor D_{OVC} ;
- G_L representa os valores limite dos grupos;
- G_C representa o número de grupos;

Início do algoritmo:

- É feita uma rotulação das regiões limiarizadas, apenas para sua identificação;
- É criado um histograma HR_n de cada região segmentada e rotulada R_n ;

- o Isto é feito usando as coordenadas obtidas na imagem rotulada, porém com os dados da imagem que contem o canal H da imagem original;
- Cada histograma HR_n é analisado, localizando-se o grupo de cores mais representativo $G_n = \{C_1, C_2, C_3, C_4, C_5\}$, onde $C_1, \dots, C_5$ são cores contíguas do vetor de histograma HR_n ;
- O grupo mais significativo é o conjunto das 5 posições contíguas de maior soma do histograma;
- É criado um vetor com os valores centrais $V_C(n)$ do grupo de cores detectado no passo 2;
 - o O valor central é o elemento C_3 do conjunto G_n calculado no passo 2;
- O vetor é ordenado numericamente gerando o vetor OV_C ;
- É criada a variável D_{ovc} contendo a derivada primeira do vetor ordenado OV_C , calculada segundo a fórmula: $D_{ovc}(e) = OV_C(e+1) - OV_C(e)$, onde $e = \{0, \dots, n-1\}$;
- É feita a média mD do vetor D_{ovc} ;
- É criado um vetor de tamanho variável chamado G_L ;
- Os valores do vetor de derivada D_{ovc} são analisados, e sempre que $D_{ovc}(e) > mD$ para $e = (0, \dots, n)$, o valor $D_{ovc}(e)$ é adicionado ao vetor G_L ;
- O número de elementos de G_L é contado para definir a variável G_C ;
- O vetor $V_C(n)$ é analisado e cada região n é classificada conforme sua distribuição entre os valores dos elementos do vetor G_L ;

Fim do algoritmo

Esta técnica dispensa o uso combinatório para solucionar o problema do agrupamento, que como visto anteriormente é extremamente custoso computacionalmente.

Para ilustrar o processo, considere a figura 8.6a já utilizada nos exemplos anteriores de morfologia, onde o número de regiões identificado foi 8 (o que pode ser verificado pela figura 8.6h), logo, “n” é igual a 8.

O conjunto das regiões segmentadas R_n é representado pelo conjunto de regiões segmentadas $R_n = \{1, 2, 3, 4, 5, 6, 7, 8\}$.

As figuras a seguir mostram os histogramas HR_n das regiões marcadas na figura 8.6h:

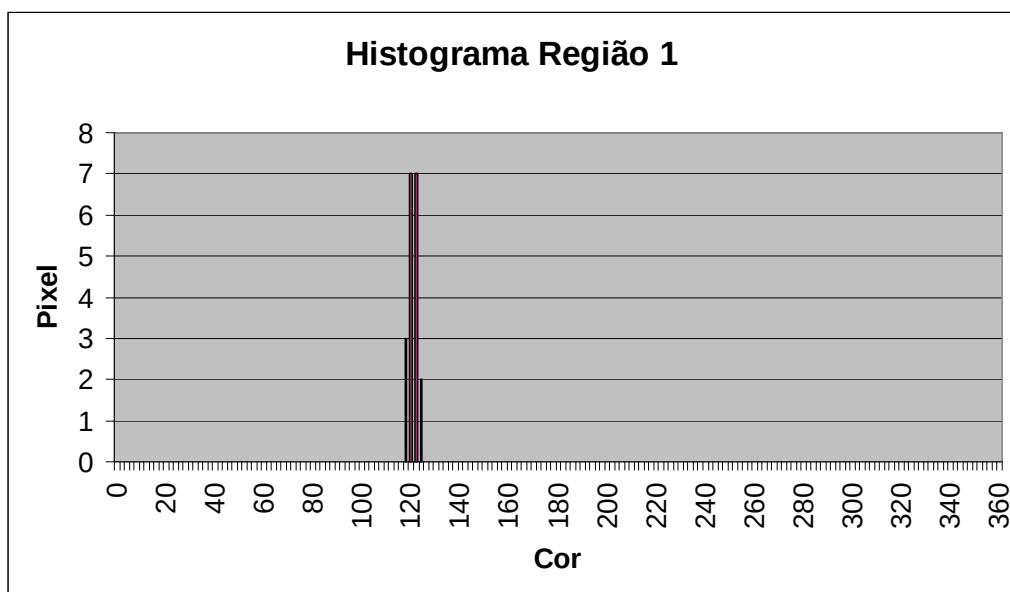


Figura 8.8 – Histograma da região segmentada 1

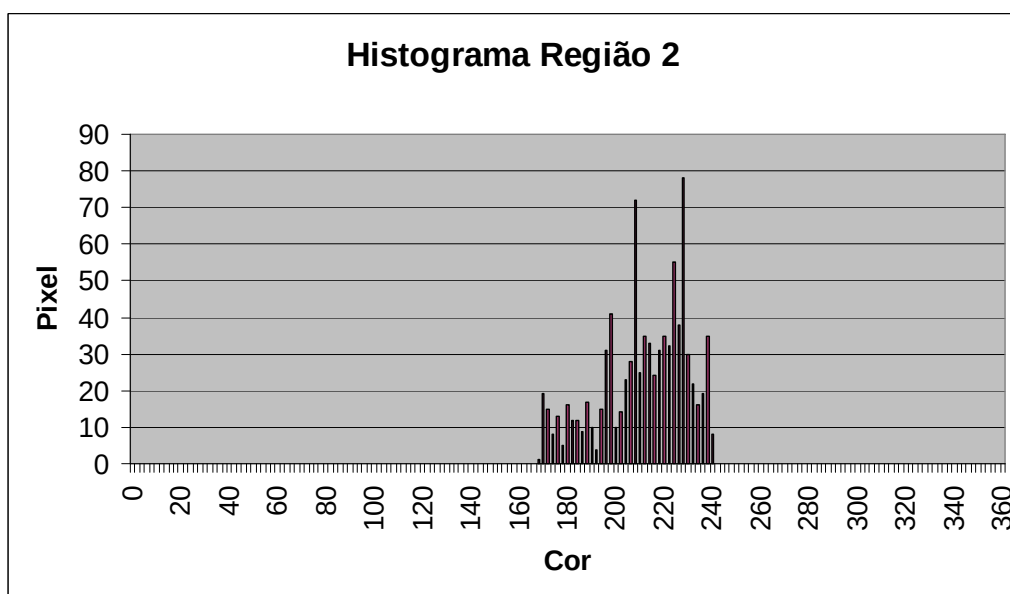


Figura 8.9 – Histograma da região segmentada 2

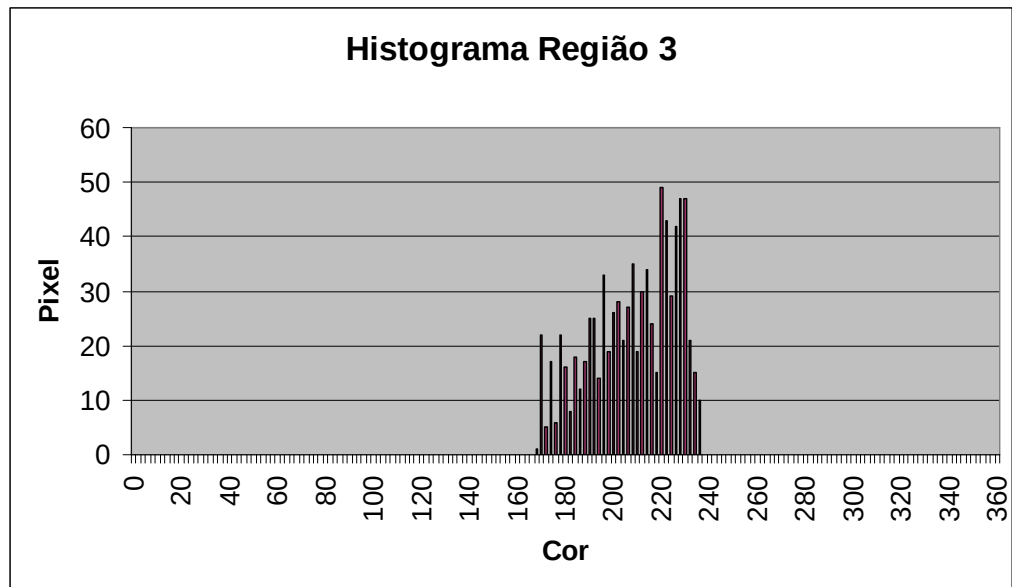


Figura 8.10 – Histograma da região segmentada 3

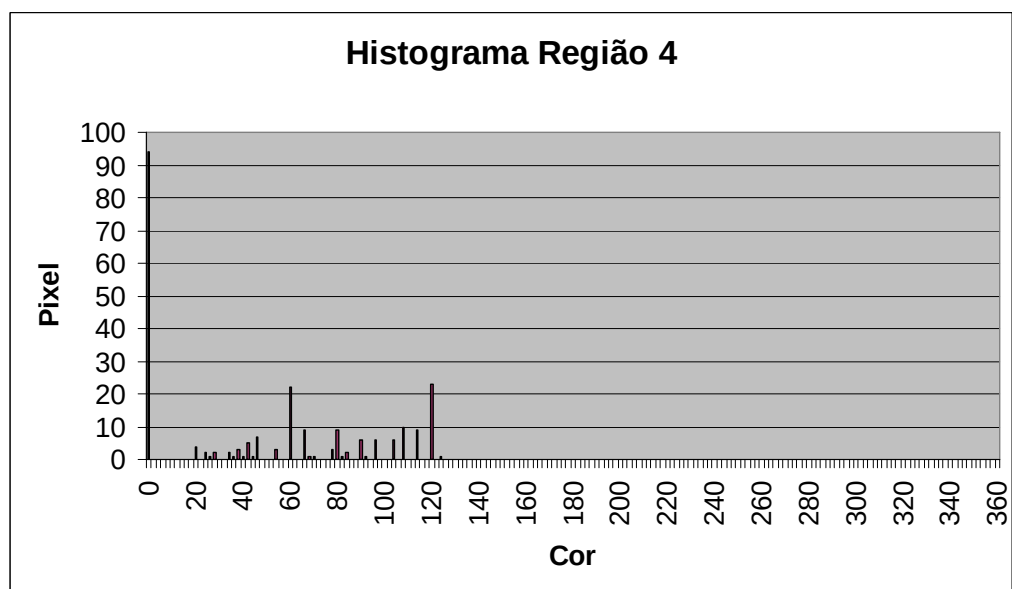


Figura 8.11 – Histograma da região segmentada 4

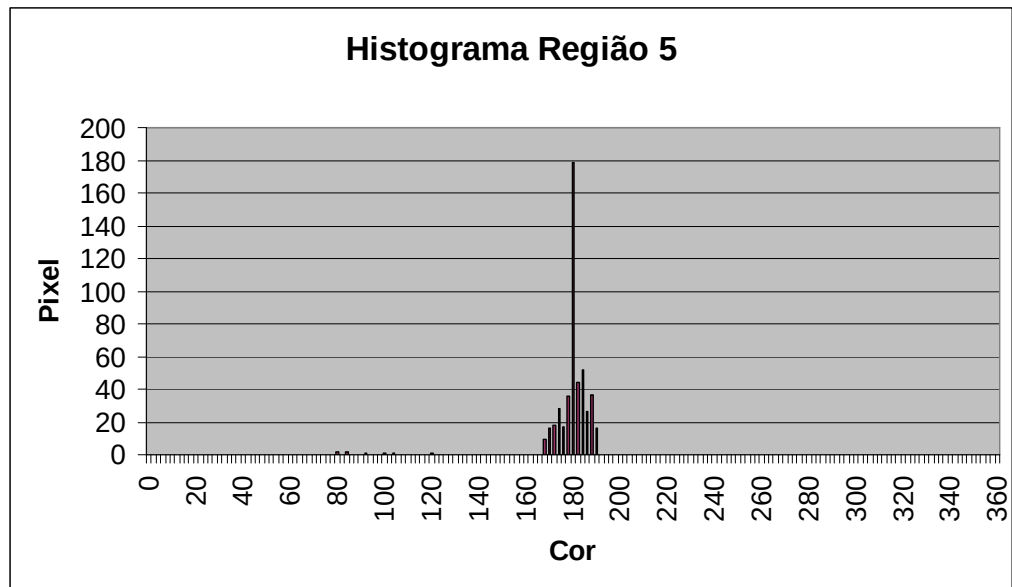


Figura 8.12 – Histograma da região segmentada 5

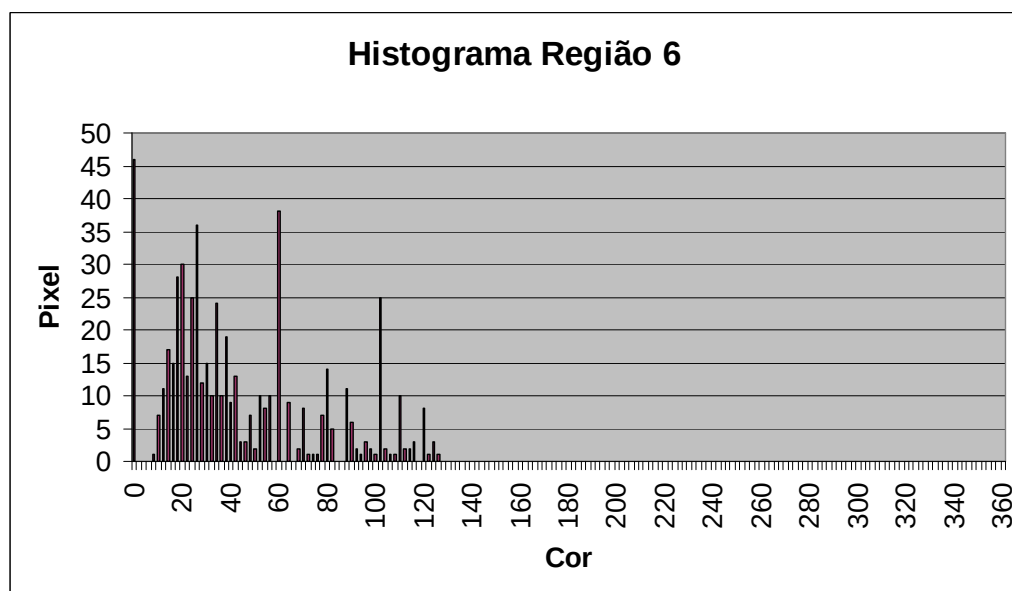


Figura 8.13 – Histograma da região segmentada 6

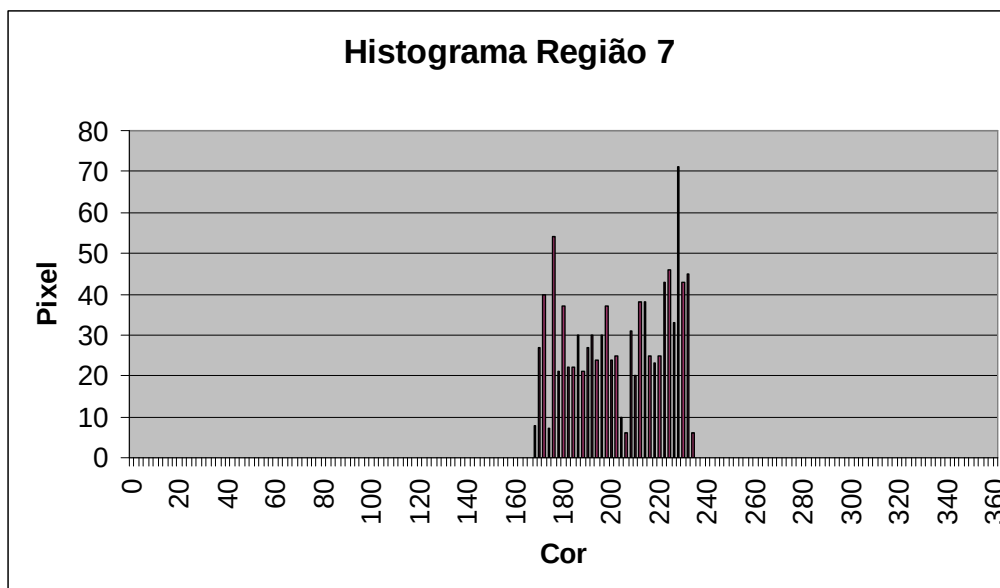


Figura 8.14 – Histograma da região segmentada 7

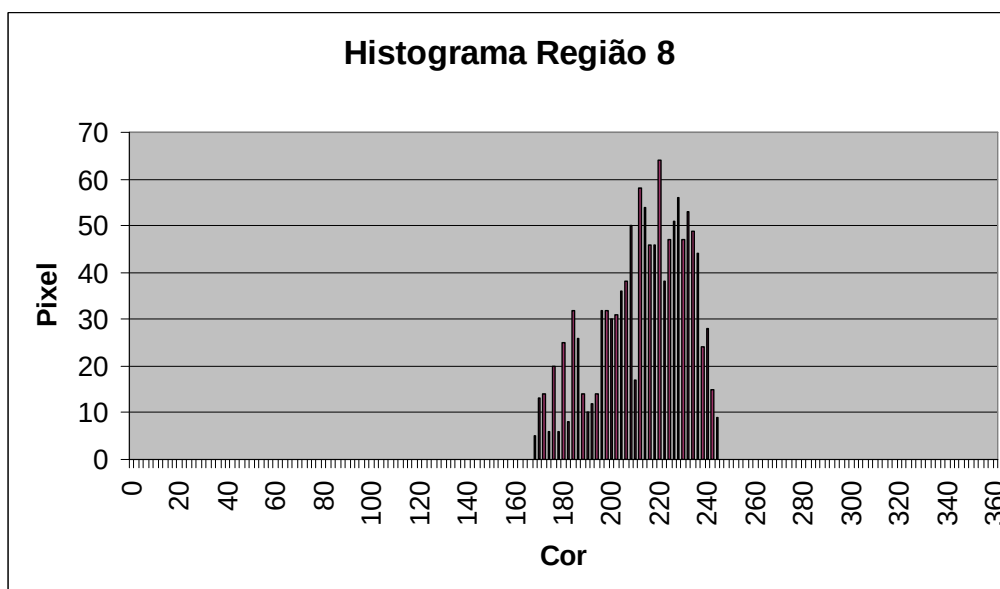


Figura 8.15 – Histograma da região segmentada 8

Os histogramas são analisados, para determinar o conjunto das 5 cores contíguas cuja soma é mais representativa, para então determinar o valor central deste conjunto de cores. Suponha o vetor $\{1,2,3,4,5,9,5,4,3,2,1\}$ de 11 posições, o maior somatório de 5 elementos contíguos se dá a partir da posição 4, ou seja, define-se o subconjunto $\{4,5,9,5,4\}$. O elemento central por definição é o elemento de posição 3 do subconjunto, ou seja, o número “9”.

Após análise, conclui-se que o conjunto V_C que tem o elemento central do grupo de cores mais representativos de cada imagem é: $V_C = \{118;224;224;4;182;22;226;216\}$.

Isto pode ser verificado pela comparação do vetor V_C com os histogramas das regiões R_n .

O conjunto V_C é ordenado dando origem ao conjunto $OV_C = \{4; 22; 118; 182; 216; 224; 224; 226\}$. Neste conjunto OV_C é aplicada a derivada primeira, gerando o conjunto $D_{OV_C} = \{0; 18; 96; 64; 34; 8; 0; 2\}$.

É então calculada a média do conjunto D_{OV_C} , dando origem ao valor $mD = 27,75$. A comparação do valor de mD com o conjunto D_{OV_C} gera “ G_C ” valores que são adicionados ao conjunto G_L . Os valores são adicionados sempre que $D_{OV_C}(e)$ for maior que mD , como pode ser observado na figura 8.16.

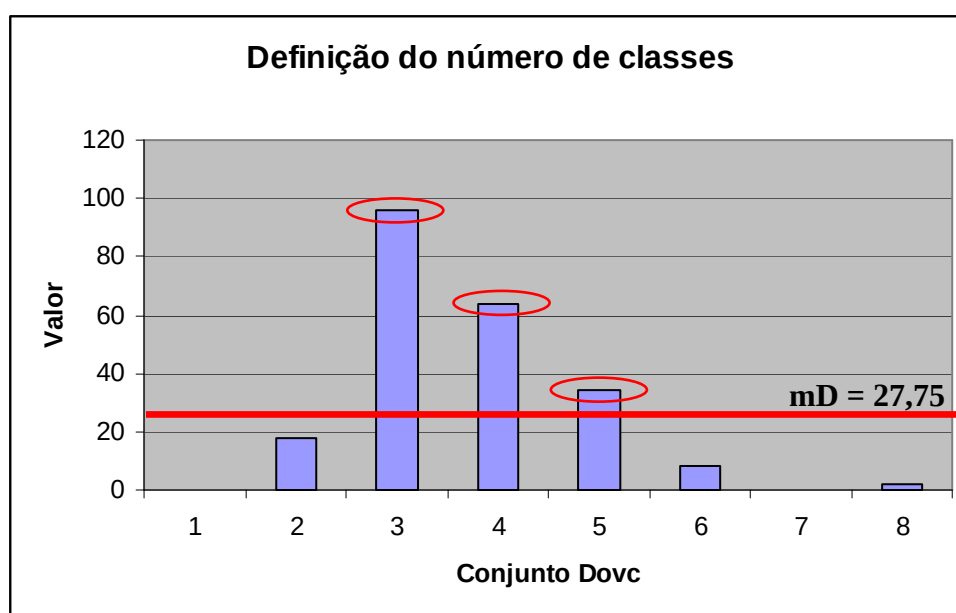


Figura 8.16 – Histograma de região segmentada

Uma vez com os valores do conjunto $G_L = \{216; 182; 118\}$, é possível saber que $G_C = 4$, e então classificar as regiões.

Sendo o conjunto $V_C = \{118; 224; 224; 4; 182; 22; 226; 216\}$, a classificação dos grupos é a seguinte:

Grupo 1: Valores de V_C menores que 118 (o menor valor de G_L)

$\{4; 22\}$ – Jogadores em branco e preto (grupo preto)

Grupo 2: Valores de V_C menores que 182 (o segundo menor valor valor de G_L)

$\{118\}$ – Juiz (grupo azul)

Grupo 3: Valores de V_C menores que 216 (o segundo menor valor valor de G_L)

$\{182\}$ – Fração não segmentada como campo (grupo vermelho)

Grupo 4: Valores de V_C maiores ou iguais a 216 {216;224;224;226} –
Jogadores em vermelho (grupo branco)

Capítulo 9

Experimentos e Discussão de Resultados

9.1. Considerações iniciais

Todos os testes foram realizados utilizando um computador Athlon XP 2600+, com 1gb de memória RAM DDR400, placa mãe Shuttle MN31/n, e sistema operacional Windows XP Professional. Para as funções primitivas de processamento de imagens, como a morfologia, carga de imagens, limiarizações, etc. foi utilizada a biblioteca de processamento de imagens OpenCV (<http://sourceforge.net/projects/opencv/>), que é de código aberto. Os programas de teste foram criados na linguagem C++, utilizando o compilador Borland C++ Builder 6.0.

9.2. Vídeos utilizados

Todos os vídeos utilizados foram digitalizados ou encontrados na internet, por esta razão, não haveria maneira de mensurar adequadamente sua qualidade (não há a captação original). A grande variedade dos vídeos encontrados também dificultaria o processo de qualificação, pois estes usam diferentes tipos de compressão e codificação, muitos deles possivelmente sofreram mais de uma etapa de compressão. Por este motivo os critérios utilizados para definir qualidade tiveram de ser abrangentes, mesmo porque servirão apenas para diferenciar os resultados dos algoritmos. Neste trabalho foram elencados três níveis de avaliação, definidos da seguinte maneira:

Ruim: Inferior à qualidade de um vídeo capturado de uma transmissão de TV, normalmente com uma compressão agressiva, ou com ruídos característicos de uma transmissão analógica.

Bom: Qualidade aproximada a de uma transmissão de TV, sem muito ruído e sem danos consideráveis causados por compressão.

Ótimo: Qualidade superior a uma transmissão de TV (qualidade de DVD, HDTV, etc.).

Atualmente não existem testes padronizados para avaliar comparativamente a eficácia dos algoritmos de detecção de região dominante, nem tampouco de segmentação de regiões. Para avaliar os algoritmos, foram providenciados 14 vídeos de futebol todos com características diferentes, e duração de pelo menos 45 minutos. Os vídeos são os seguintes (Tabela 9.1):

Vídeo:	Codec:	Bitrate:	FPS:	Resolução:	Proporção x:y:	Qualidade:
1	XviD	1074	25.000	352x288	(1.22:1) [=11:9]	RUIM
2	XviD	801	25.000	512x304	(1.68:1) [=32:19]	RUIM
3	DivX 5.0	1507	25.057	512x380	(1.35:1) [~31:23]	BOM
4	XviD	1825	25.000	512x384	(1.33:1) [=4:3]	BOM
5	XviD	782	25.000	576x320	(1.80:1) [=9:5]	BOM
6	XviD	866	25.000	512x384	(1.33:1) [=4:3]	BOM
7	XviD	948	25.000	640x480	(1.33:1) [=4:3]	BOM
8	XviD	986	25.000	640x368	(1.74:1) [=40:23]	BOM
9	DivX 5.0	864	25.000	640x480	(1.33:1) [=4:3]	RUIM
10	XviD	1653	29.971	672x368	(1.83:1) [=42:23]	ÓTIMO
11	XviD	2349	29.970	960x544	(1.76:1) [=30:17]	ÓTIMO
12	XviD	2397	29.970	960x544	(1.76:1) [=30:17]	ÓTIMO
13	XviD	2704	29.970	960x544	(1.76:1) [=30:17]	ÓTIMO
14	XviD	2604	29.970	960x544	(1.76:1) [=30:17]	ÓTIMO

Tabela 9.1 – Descrição dos vídeos usados

Os nomes dos vídeos da tabela 9.1 podem ser observados na tabela 9.2.

Vídeo:	Nome do vídeo:
1	Confeds Cup Final Brazil vs Argentina 1st Half.avi
2	germany vs brazil 2nd half.avi
3	Final.Copa.del.Rey.2006.Espanyol-Zaragoza.HQTVRip.avi
4	FIFA.World.Cup.2006.Espanya.Vs.Tunez CD2.avi
5	gsf-wm.2006.england.vs.paraguay.avi
6	gsf-wm.2006-england.vs.ecuador.avi
7	2006-01-02 - epl - newcastle_vs_middlesbrough.avi
8	UEFA Cup SF 2nd L - Boro v Bucharest - 2nd Half.avi
9	WC98 Brasil-France.avi
10	Copa2006.BRASILxAUSTRALIA.1oTempo.Portugues.xara.avi
11	World.Cup.2006.Group.A.Germany.vs.Costa.Rica.HR.HDTV.avi
12	World.Cup.2006.Group.A.Poland.vs.Ecuador.avi
13	World.Cup.2006.Group.B.Trinidad.and.Tobago.vs.Sweden.HR.HDTV.avi
14	world.cup.2006.group.d.iran.vs.mexico.hr.hdtv.ac3.5.1.xvid.cd2.avi

Tabela 9.2 – Legenda dos nomes dos vídeos

Alguns vídeos apesar de serem comprimidos com um algoritmo que preserva suas características, a fonte do qual foram obtidos nem sempre é de boa qualidade. Em alguns casos o vídeo foi comprimido e depois convertido para outro tipo de compressão, comprometendo ainda mais a qualidade.

A comparação mesmo que superficial com outros métodos existentes também não foi possível, pois não foram encontrados autores que disponibilizam o material usado para teste. Tendo em vista estas dificuldades, é parte deste trabalho disponibilizar na internet os trechos de vídeo utilizados bem como os resultados obtidos para que outros autores tenham condição de avaliar comparativamente seus trabalhos, contribuindo para a evolução nesta área de pesquisa.

9.3. Qualidade

A avaliação da qualidade da segmentação exigiu a criação de uma métrica de qualidade, que consiste em comparar o resultado desejado com o obtido, dando pesos para as diversas partes do processamento. São eles:

1. Região predominante na imagem
2. Região adjacente na imagem (região que não faz parte do campo nem do conjunto de jogadores, ex: arquibancadas, letreiros de publicidade, etc.)
3. Jogadores/Juiz segmentados corretamente na imagem
4. Jogadores/Juiz classificados corretamente na imagem
5. Elementos segmentados e classificados corretamente na imagem

O cálculo dos itens 1 e 2 são feitos pixel a pixel, e dos itens 3, 4, e 5 por meio de inspeção visual das regiões. Cada item vai gerar um percentual de qualidade de até 100%, que será multiplicado por um peso, e depois todos os percentuais são somados, totalizando um percentual de qualidade para a imagem processada.

Neste trabalho, os itens 1 e 2 (tabela 9.3) contam com um peso maior, pois representam o processo principal, do qual dependem todos os outros, e que, caso não seja bem feito, todos os outros resultarão errados. Existe também o fato deste processo ter de classificar todos os pixels da imagem, enquanto os outros processos tratam de classificar apenas uma pequena porção. Foram feitos experimentos para avaliar os resultados com o uso de diferentes pesos, e no apêndice C são mostradas tabelas que ilustram a variação dos índices de qualidade para as várias combinações, mostrando que estas mudanças dos pesos não influenciam significativamente o bom resultado atingido pela técnica.

Ítem	Cálculo	Peso (%)
1	Pixels corretos de campo / Total de pixels de campo	35
2	Pixels corretos de regiões adj. / Total de pixels de regiões adj.	35
3	Jogadores segmentados corretamente / Total de jogadores	10
4	Jogadores classificados corretamente / Total de jogadores	10
5	Elementos classificados corretamente / Elementos segmentados	10

Tabela 9.3 – Critérios de medição de qualidade

A classificação manual das regiões é procedimento baseado no julgamento subjetivo de um operador, logo existe uma margem de erro na medição destas características. Para minimizar este erro, as marcações foram feitas por dois operadores diferentes, e depois conferidas com o auxílio de um programa desenvolvido especialmente para isto. Um dos operadores tem grande experiência em aquisição e processamento de imagens, e o outro é considerado leigo no assunto.

Nas imagens onde a diferença absoluta das marcações foi superior a 5%, o procedimento foi refeito. A diferença das marcações foi calculada segundo a fórmula 9.1.

$$D = \left(\frac{Dp}{A \times L} \right) \times 100 \quad (9.1)$$

Onde:

D = Diferença percentual das imagens

A = Altura em pixels da imagem

L = Largura em pixels da imagem

Dp = Pixels de mesma posição, mas cores diferentes entre as duas imagens

O cálculo é realizado comparando pixel a pixel a imagem de marcação manual e a de marcação automática.

O processo de classificação é descrito a seguir:

1. Foram capturadas aleatoriamente um mínimo de 20 imagens de cada um dos 14 vídeos disponíveis;
2. As imagens que não continham campo, cujo zoom era muito grande, ou que de alguma forma fugiam do escopo desta aplicação foram eliminadas, e então foram tomadas novas imagens para substituir as descartadas, exemplos de imagens não aproveitadas podem ser vistos na figura 9.10;
3. As imagens aproveitadas foram abertas em um software de edição de imagens (Paint .Net v2.63, <http://www.eecs.wsu.edu/paint.net/>)
4. As regiões de campo, jogadores e regiões adjacentes foram marcadas conforme a seguinte tabela:

Elemento	Cor
Campo	Verde
Regiões adjacentes (arquibancadas, placares, etc.)	Vermelho
Jogadores time 1	Amarelo
Jogadores time 2	Ciano
Juiz/Goleiro	Roxo
Outros elementos em campo	Roxo

Tabela 9.4 – Descrição da marcação manual das imagens

5. As imagens marcadas manualmente pelos operadores (figura 9.1) foram comparadas, e quando apresentavam mais de 5% de discrepância, eram corrigidas pelo operador mais experiente;
6. O sistema gerou imagens com o mesmo tipo de marcação automaticamente;
7. As diferenças entre as imagens geradas manual e automaticamente foram calculadas através de comparação direta entre os pixels;
8. As informações sobre os jogadores e outros elementos em campo foram comparadas manualmente, verificando se estavam sobrepostas;
9. Foi gerada, então, uma tabela com os dados obtidos;

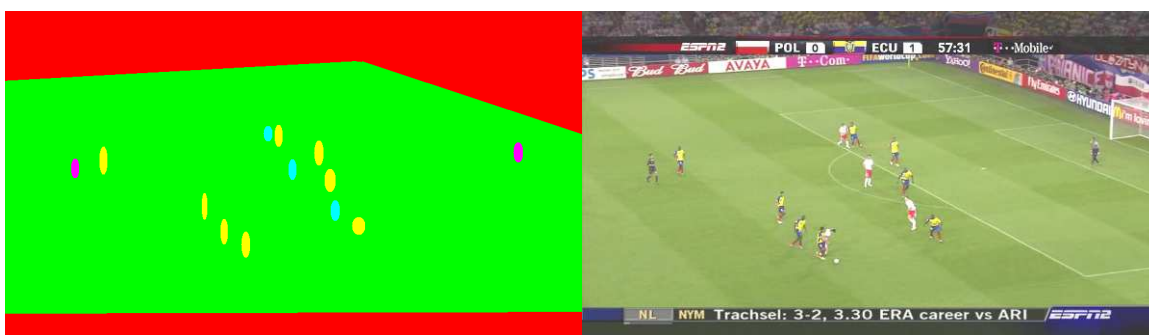


Figura 9.1 – (a) Exemplo de marcação manual de imagem, (b) Imagem Original

No item 8, a comparação não deve ser absoluta, pois para a detecção dos jogadores as imagens passam por operações morfológicas, e somado a isto, há mistura de cores nas regiões de fronteira dos objetos, por isto não faz sentido uma comparação volumétrica da região dos jogadores e sim uma correta detecção do seu posicionamento. Marcações considerando o volume seriam desejáveis apenas nos casos onde a precisão fosse crítica, porém existem técnicas diferentes para este tipo de problema, normalmente envolvendo detecção de borda, textura, etc. Marcações não centralizadas ou inconsistentes foram consideradas incorretas. Exemplos dos critérios de marcação podem ser vistos na figura 9.2, onde as marcações em branco foram consideradas corretas e em vermelho erradas.



Figura 9.2 – Exemplo de marcação correta (branco) e incorreta (vermelho).

A classificação considerada correta é a que separa elementos distintos, como por exemplo, jogadores de cada time de uma cor, juiz de outra cor, goleiro de uma terceira. No caso de elementos de mesma cor, como juiz e goleiro em preto, por exemplo, é considerada correta a classificação no mesmo grupo. O mesmo acontece com enfermeiros atendendo em campo, que terão a mesma classificação que um time que usa branco.

Eventualmente aparecem porções de campo que pelo fato de terem cores diferentes da predominante no resto do campo, são marcadas como algum elemento distinto. Nestes casos será considerada uma marcação correta se estes elementos distintos estiverem classificados em grupos diferentes dos jogadores, juiz e dos outros elementos semanticamente diferentes. Estas regiões normalmente chamadas de “ruídos”, ou “lixo” na imagem têm inúmeras explicações, que vão desde uma região com barro no gramado até uma distorção causada na compressão. Com os recursos atuais não é possível detectar e eliminar todas estas regiões, especialmente sem o uso de heurísticas, por este motivo a solução encontrada é separá-las dos outros objetos semanticamente relevantes, para que em outro momento posterior possam ser eliminadas. Este processo de identificação de regiões não faz parte do escopo deste trabalho, pois se trata de uma operação de mais alto nível, e dependente de aplicação.

Nos casos de oclusão ou fusão de jogadores, como mostrado na figura 9.3, é marcado apenas o jogador cuja cor de camisa é predominante em termos espaciais. Pelo

fato dos algoritmos usados não serem feitos para detectar oclusão nem fusão, e ainda em virtude de a classificação dos jogadores ser feita pela cor da camisa, a resposta esperada do sistema é exatamente considerar jogadores fundidos como um, e classificá-lo como o da cor predominante.

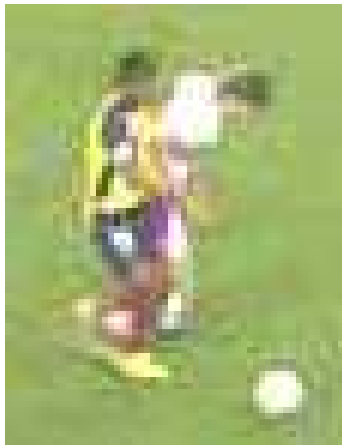


Figura 9.3 – Oclusão/fusão de jogador

Para contabilizar as regiões marcadas corretamente foi usado o critério da maioria, ou seja, supondo cinco (5) regiões vermelhas, das quais quatro (4) são pedaços de campo, e uma é um jogador, são consideradas quatro (4) regiões marcadas corretamente. Supondo que haja ainda mais três (3) jogadores do mesmo time, mas marcados com azul, são apenas três (3) classificados corretamente, pois o marcado em vermelho é considerado uma marcação errada.

Os procedimentos de inspeção visual para o cálculo dos itens 3, 4, e 5 são descritos a seguir:

1. O total do número de regiões marcadas é fornecido automaticamente no processamento das imagens;
2. São contabilizadas visualmente as marcações das regiões corretamente classificadas (cores iguais para grupos semanticamente semelhantes);
3. É contado o número total de jogadores, incluindo os que estão parcialmente na imagem e os que estão nas extremidades;
4. São contados os jogadores que estão ocultos ou na extremidade da imagem;
5. São contados os jogadores classificados corretamente
6. São contados os jogadores que não estão nas extremidades e que não foram segmentados corretamente;

A figura 9.4d ilustra a diferença entre uma marcação manual e a marcação automática. As regiões de cor verde escuro são aquelas onde a marcação automática de campo coincidiu com a manual, e as marcações em verde claro são aquelas onde houve diferença. O mesmo ocorre para as marcações de cor vermelho escuro, que representam os lugares onde coincidiram as marcações manual e automática para as regiões adjacentes, e as vermelho escuro mostram as diferenças. As regiões que estão marcadas em vermelho claro, dentro da grande região vermelho escuro, são áreas que foram classificadas como campo. Isto ocorre pelo fato de existirem grupos de pixels da mesma cor do campo nas regiões adjacentes.

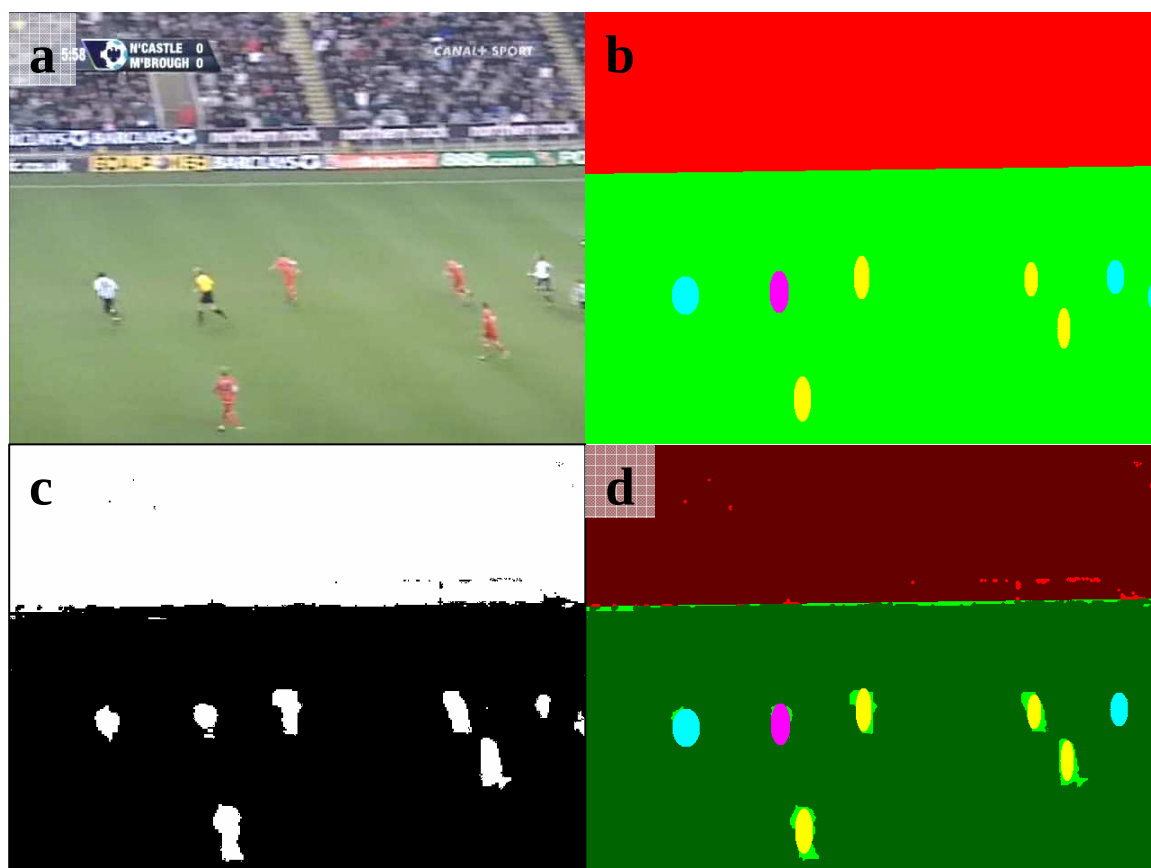


Figura 9.4 – (a) é a imagem original, (b) a imagem de marcação manual, (c) a imagem limiarizada, e (d) representa a diferença entre a marcação manual e a automática.

9.4. Resultados

Os resultados dos testes serão demonstrados por meio de dois tipos de planilha, sendo a primeira com os dados coletados manual e automaticamente, e a segunda, com estes dados compilados e mostrados de maneira percentual. Serão chamadas de

“planilha de dados” e “planilha de resultados” respectivamente. Os valores de ambas as planilhas estão discriminados por quadro, e no final, uma média de todos os quadros.

Legendas da planilha de dados:

CIM – Total de pixels de campo classificados corretamente

CIA – Total de pixels de campo (Imagem Manual)

RAIM – Total de pixels de região adjacente classificados corretamente

RAIA – Total de pixels de região adjacente campo (Imagem Manual)

TRS – Total de regiões segmentadas

RCC – Regiões classificadas corretamente

TJ – Número total de jogadores (entenda-se jogadores todos os elementos em campo, incluindo juiz, bandeirinha, etc.) visíveis na imagem

JCA – Número total de jogadores (entenda-se jogadores todos os elementos em campo, incluindo juiz, bandeirinha, etc.) colados nas regiões adjacentes ou nas extremidades da imagem

JCC – Número total de jogadores classificados corretamente

JNS – Número total de jogadores não segmentados

Legendas da planilha de resultados:

CMPO – Percentual de acerto na classificação de pixels de campo

ADJ – Percentual de acerto na classificação de pixels de regiões adjacentes

JSC1 – Jogadores segmentados corretamente

JSC2 – Jogadores segmentados corretamente*

JCC1 – Jogadores classificados corretamente

JCC2 – Jogadores classificados corretamente*

RCCP – Regiões classificadas corretamente

IQ1 – Índice de qualidade

IQ2 – Índice de qualidade*

* Desconsiderando jogadores colados nas bordas

Fórmulas utilizadas na planilha de resultados:

$CMPO = CIM / CIA$

$ADJ = RAIM / RAIA$

$JSC1 = (TJ - JCA - JNS) / TJ$

$$JSC2 = (TJ - JCA - JNS) / (TJ - JCA)$$

$$JCC1 = JCC / (TJ - JNS)$$

$$JCC2 = JCC / (TJ - JNS - JCA)$$

$$RCCP = RCC / TRS$$

$$IQ1 = (CMPO * 0,35 + ADJ * 0,35 + JSC1 * 0,10 + JCC1 * 0,10 + RCC * 0,10)$$

$$IQ2 = (CMPO * 0,35 + ADJ * 0,35 + JSC2 * 0,10 + JCC2 * 0,10 + RCC * 0,10)$$



Figura 9.5 – Exemplo de imagem já classificada

Para exemplificar o processo de cálculo, considerando a figura 9.5:

A tabela de dados ficaria da seguinte forma:

CIM	CIA	RAIM	RAIA	TRS	RCC	TJ	JCA	JCC	JNS
186683	189643	111665	111665	8	8	8	1	7	0

Tabela 9.5 – Exemplo de tabela de dados

Onde:

CIM – Total de pixels de campo classificados corretamente é 186683

CIA – Total de pixels de campo é 189643

RAIM – Total de pixels de região adjacente classificados corretamente é 111665

RAIA – Total de pixels de região adjacente é 111665

TRS – Total de regiões segmentadas é 8 (quatro quadrados brancos, dois pretos, um azul e um vermelho)

RCC – Regiões classificadas corretamente (todas foram corretamente segmentadas, pois o time de camisa vermelha está no grupo branco, o de camisa e calção preto no grupo preto, o juiz no grupo azul e uma fração de campo está isolada no grupo vermelho)

TJ – Número total de jogadores visíveis na imagem (são ao total oito, sendo sete marcados e mais um aparecendo parcialmente na extremidade direita da imagem)

JCA – Número total de jogadores colados nas regiões adjacentes ou nas extremidades (apenas um, aparecendo parcialmente na extremidade direita da imagem)

JCC – Sete jogadores classificados corretamente (marcados nas cores branco, preto e azul)

JNS – Nenhum jogador não segmentado (aqui se contabilizam os que não estão colados nas extremidades, que são contabilizados no parâmetro JCA)

A partir destes dados é calculada a tabela de resultados:

CMPO	ADJ	JSC1	JSC2	JCC1	JCC2	RCCP	IQ1	IQ2
98,44%	100,00%	87,50%	100,00%	87,50%	100,00%	100,00%	96,95%	99,45%

Tabela 9.6 – Exemplo de tabela de resultados

Onde:

CMPO – 98,44% é o percentual de acerto na classificação de pixels de campo, pois 98,44% dos pixels classificados automaticamente estavam em regiões marcadas manualmente como campo.

ADJ – 100,00% é o percentual de acerto na classificação de pixels de regiões adjacentes, pois 100,00% dos pixels classificados automaticamente estavam em regiões marcadas manualmente como região adjacente.

JSC1 – 87,50% pois dos 8 jogadores da imagem 7 foram classificados corretamente

JSC2 – 100,00% pois se forem excluídos os jogadores fundidos às extremidades do campo e da imagem, os 7 jogadores enquadrados nesta condição foram classificados corretamente

JCC1 – 87,50% pois dos oito jogadores presentes na imagem 7 foram classificados corretamente

JCC2 – 100% pois dos todos os sete jogadores não fundidos nas extremidades foram classificados corretamente

RCCP – 100,00% pois todas as oito regiões segmentadas foram separadas em grupos semanticamente coerentes (jogadores do mesmo time, juiz, ruído)

IQ1 – 96,95% segundo a fórmula ($CMPO * 0,35 + ADJ * 0,35 + JSC1 * 0,10 + JCC1 * 0,10 + RCC * 0,10$) que aplica os pesos para cada característica

IQ2 – 99,45% segundo a fórmula ($CMPO * 0,35 + ADJ * 0,35 + JSC2 * 0,10 + JCC2 * 0,10 + RCC * 0,10$) que aplica os pesos para cada característica, neste caso, excluindo as regiões fundidas nas extremidades.



Figura 9.6 – Exemplo de imagem já classificada

Exemplificando o processo de cálculo, considerando a figura 9.6:

A tabela de dados ficaria da seguinte forma:

CIM	CIA	RAIM	RAIA	TRS	RCC	TJ	JCA	JCC	JNS
443845	447818	63296	69739	11	10	12	0	10	1

Tabela 9.7 – Exemplo de tabela de dados

Onde:

CIM – Total de pixels de campo classificados corretamente é 443845

CIA – Total de pixels de campo é 447818

RAIM – Total de pixels de região adjacente classificados corretamente é 63296

RAIA – Total de pixels de região adjacente é 69739

TRS – Total de regiões segmentadas é 11 (cinco vermelhas, quatro brancas, uma preta e uma azul)

RCC – Regiões classificadas corretamente (dez, pois das onze classificadas uma, que é a região azul, deveria estar no grupo das regiões brancas)

TJ – Número total de jogadores visíveis na imagem (são ao total doze, todos bem visíveis na imagem)

JCA – Não há nenhum jogador colado nas regiões adjacentes ou nas extremidades

JCC – Dez jogadores foram classificados corretamente, o décimo primeiro jogador, em azul, foi classificado em um grupo distinto dos outros de mesmo time.

JNS – Um jogador não foi segmentado (em amarelo, sob a linha do campo)

A partir destes dados é calculada a tabela de resultados:

CMPO	ADJ	JSC1	JSC2	JCC1	JCC2	RCCP	IQ1	IQ2
99,11%	90,76%	91,67%	91,67%	90,91%	90,91%	90,91%	93,80%	93,80%

Tabela 9.8 – Exemplo de tabela de resultados

Onde:

CMPO – 99,11% é o percentual de acerto na classificação de pixels de campo, pois 99,11% dos pixels classificados automaticamente estavam em regiões marcadas manualmente como campo.

ADJ – 90,76% é o percentual de acerto na classificação de pixels de regiões adjacentes, pois 90,76% dos pixels classificados automaticamente estavam em regiões marcadas manualmente como região adjacente.

JSC1 – 91,67% pois dos 12 jogadores da imagem 11 foram segmentados corretamente

JSC2 – 91,67% pois se forem excluídos os jogadores fundidos às extremidades do campo e da imagem (nenhum), dos 12 enquadrados nesta condição, 11 foram classificados corretamente

JCC1 – 90,91% pois dos 12 jogadores presentes na imagem 10 foram classificados corretamente

JCC2 – 90,91% pois dos 12 jogadores presentes na imagem 10 foram classificados corretamente, não havendo nenhum jogador fundido com as extremidades

RCCP – 90,91% pois dentre as 12 regiões segmentadas, 11 foram separadas em grupos semanticamente coerentes (o jogador marcado em azul não estava no grupo branco)

IQ1 – 93,80% segundo a fórmula ($CMPO * 0,35 + ADJ * 0,35 + JSC1 * 0,10 + JCC1 * 0,10 + RCC * 0,10$) que aplica os pesos para cada característica

IQ2 – 93,80% segundo a fórmula ($CMPO * 0,35 + ADJ * 0,35 + JSC2 * 0,10 + JCC2 * 0,10 + RCC * 0,10$) que aplica os pesos para cada característica, excluindo as regiões fundidas nas extremidades (neste caso nenhuma)

As planilhas apresentando os resultados obtidos, bem como exemplos de imagens dos respectivos testes, podem ser verificadas no apêndice A, e no apêndice B vários quadros processados, mostrando o quadro original, bem como os intermediários e o resultado final da classificação. No apêndice C são demonstradas tabelas de resultado com diferentes pesos para a geração dos índices de qualidade, priorizando de maneira diferente as partes distintas da metodologia.

A tabela 9.9 apresenta de maneira compacta o resultado final dos 14 vídeos analisados. As duas últimas colunas são dos índices gerais de qualidade, e demonstram a eficiência dos algoritmos desenvolvidos. As tabelas 9.1 e 9.2 trazem mais informações sobre os vídeos.

Vídeo	CMPO	ADJ	JSC1	JSC2	JCC1	JCC2	RCCP	IQ1	IQ2
1	98,32%	80,75%	76,92%	82,47%	58,62%	63,75%	64,06%	82,64%	83,70%
2	78,69%	87,43%	58,20%	68,75%	66,91%	84,55%	82,26%	78,88%	81,70%
3	97,85%	98,56%	78,69%	98,46%	51,04%	64,06%	61,79%	87,90%	91,18%
4	98,87%	69,34%	90,67%	98,87%	46,60%	50,86%	71,13%	79,71%	80,96%
5	99,42%	95,53%	65,00%	69,40%	57,48%	63,08%	64,23%	86,90%	87,90%
6	98,67%	92,10%	91,79%	100,00%	58,97%	64,25%	66,12%	88,46%	89,81%
7	98,10%	98,93%	77,86%	100,00%	58,57%	75,23%	80,65%	90,67%	94,55%
8	98,76%	95,27%	85,65%	96,95%	66,36%	75,39%	71,91%	90,30%	92,34%
9	97,62%	87,62%	51,67%	59,24%	44,83%	55,91%	73,54%	81,84%	83,70%
10	95,93%	78,48%	80,73%	84,62%	61,83%	65,34%	65,89%	81,89%	82,63%
11	98,72%	92,63%	82,05%	87,07%	72,63%	77,73%	74,82%	89,92%	90,94%
12	98,83%	77,13%	71,49%	83,58%	38,12%	45,83%	68,04%	79,35%	81,33%
13	95,53%	94,73%	85,36%	95,33%	62,88%	70,59%	83,75%	89,79%	91,56%
14	63,48%	87,50%	72,80%	79,50%	66,51%	74,21%	77,10%	74,48%	75,92%

Tabela 9.9 – Resultado da análise dos vídeos

9.5. Tempo de Processamento

Durante o período de testes, foi observado que o tempo de processamento é linear, sendo uma relação direta do tamanho da imagem. No gráfico a seguir, foi traçada uma relação entre o tempo de processamento e o número de pixels. A linearidade do aumento do tempo de processamento pelo tamanho da imagem pode ser claramente observada.

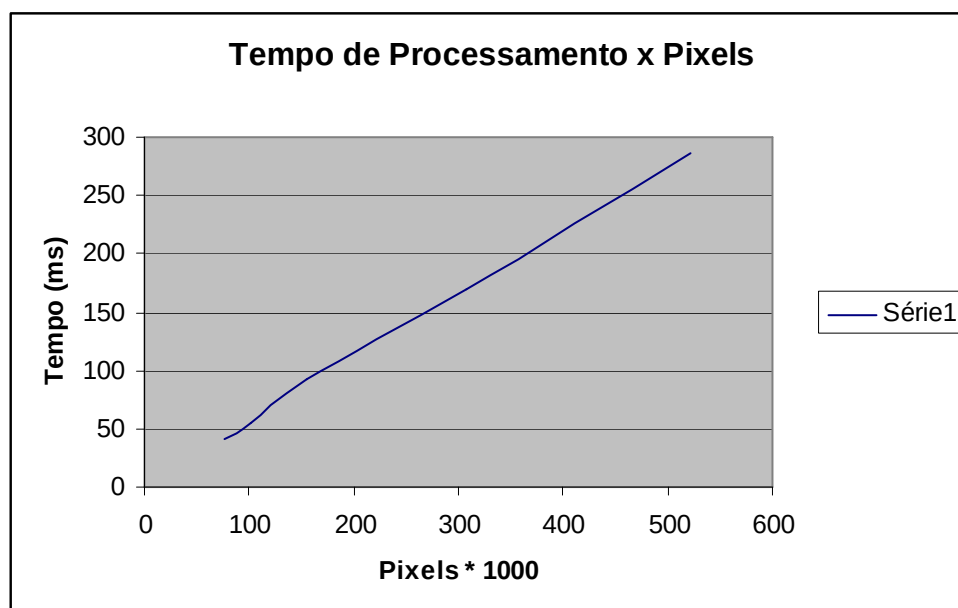


Figura 9.7 – Gráfico de Tempo de Processamento x Pixels

Foi traçado também um gráfico relacionando o tempo de processamento por pixel com o número de pixels da imagem, provando que a complexidade e a qualidade da imagem não influenciam substancialmente o tempo de processamento por pixel. Esta observação é importante, pois normalmente nos jogos de mais alta resolução, as cenas mostram uma parte maior do campo, onde são vistos mais jogadores, e mais detalhes em geral. Isto implica em um maior número de regiões segmentadas.

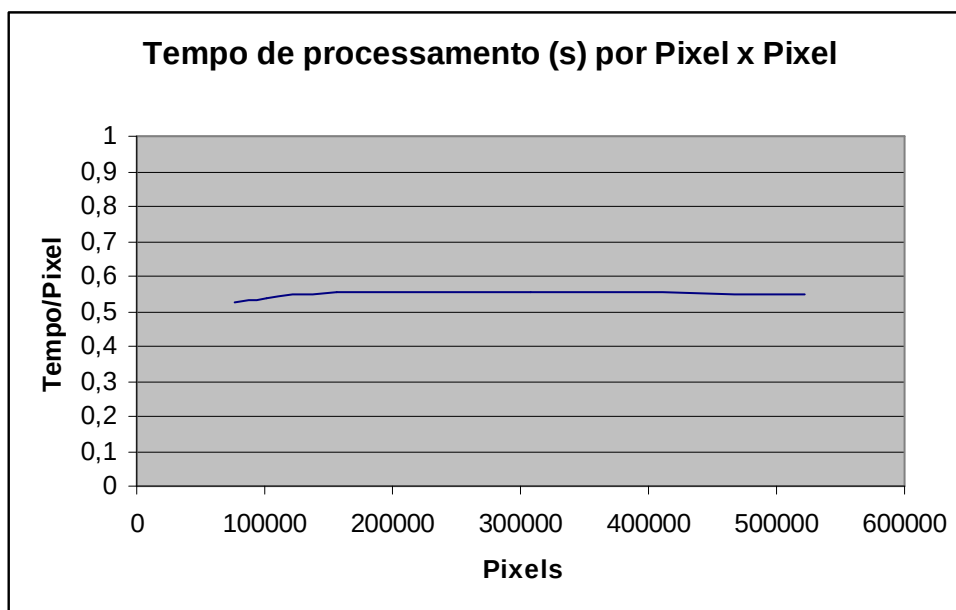


Figura 9.8 - Gráfico de Tempo de Processamento (s) por Pixel x Pixel

Uma análise mais detalhada mostrou, conforme verificado pelo gráfico da figura 9.9, que a maior parte do tempo de processamento é consumida pelas operações morfológicas. O motivo pelo qual isto ocorre é pela quantidade de informação processada por cada operação morfológica, que envolve sempre mais de uma imagem, e processamento pixel a pixel. Logo a seguir vem o tempo de conversão para HSV, que também executa operações envolvendo todos os pixels da imagem. A segmentação do campo envolve toda a imagem, porém apenas para a geração do histograma da imagem H, que é um processo rápido, e o processo de classificação das regiões é pouco custoso computacionalmente, pois envolve apenas operações elementares e em pequena quantidade.

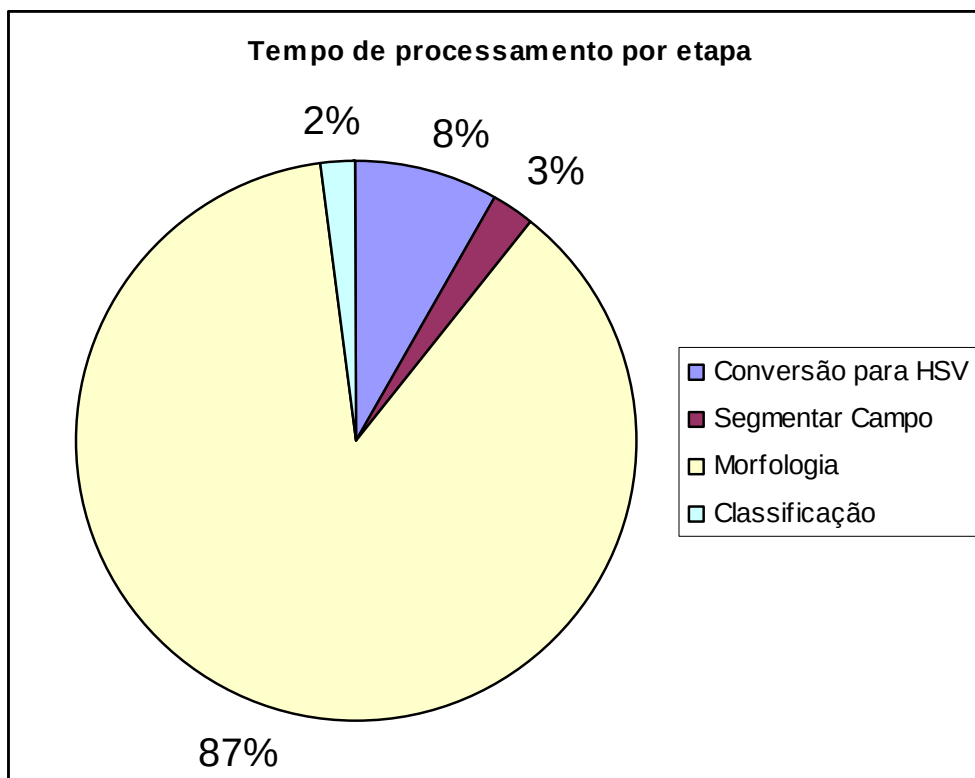


Figura 9.9 – Gráfico de Tempo de Processamento por Etapa

No equipamento utilizado, o tempo médio de processamento de cada quadro nas diversas resoluções pode ser verificado na tabela 9.10.

Resolução	Tempo (ms)
320x240	40
352x288	54
512x304	93
640x480	170
960x544	285

Tabela 9.10 - Tempo médio de processamento para diferentes resoluções de imagem

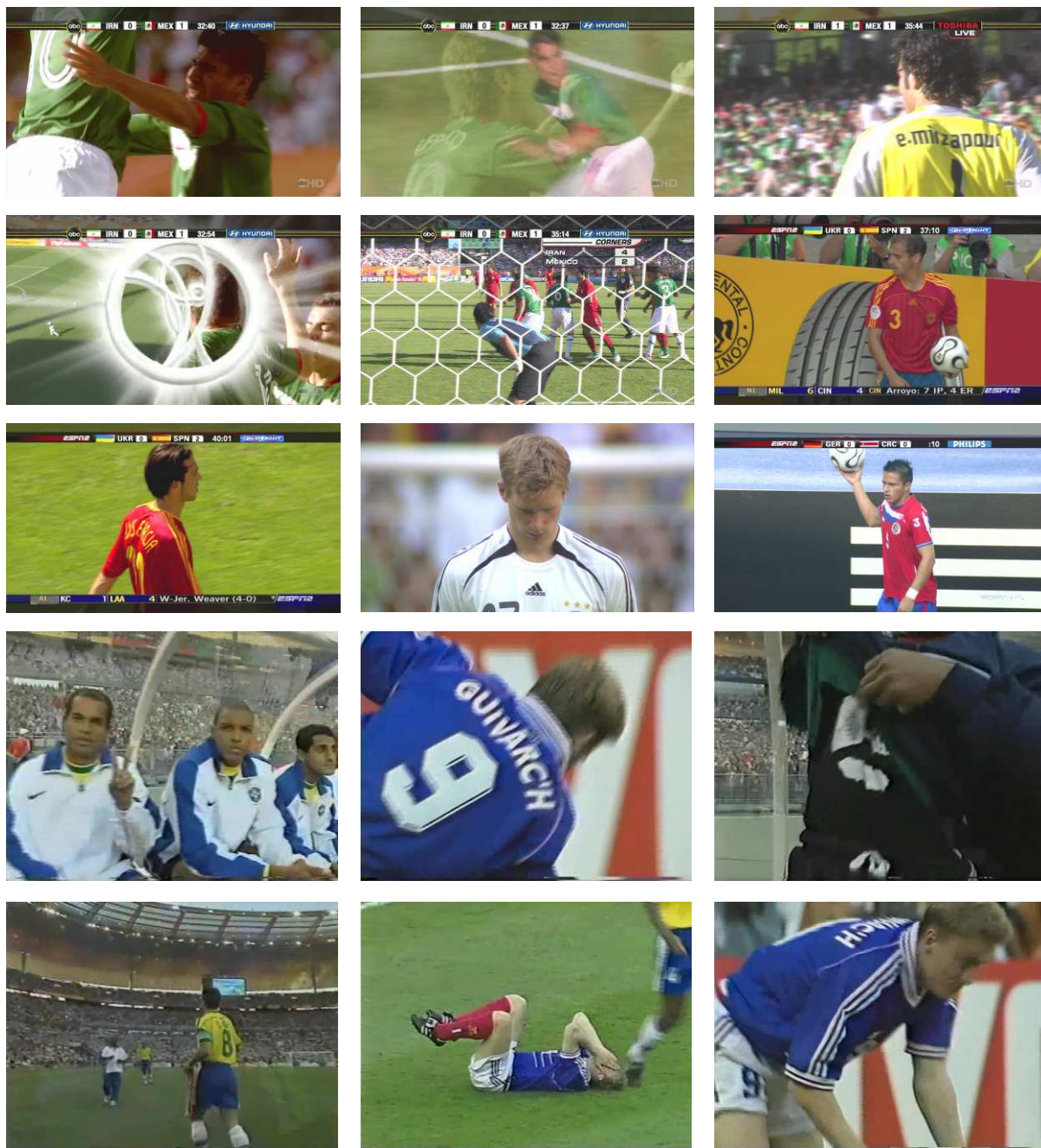


Figura 9.10 – Exemplos de imagens não aproveitadas

9.6. Problemas encontrados

Uma das maiores dificuldades encontradas, como já era esperado, foi a grande diferença de tonalidades dentro do campo. Um fenômeno comum na captação de imagens é a saturação de uma região da imagem, e isto se deve a diferença de iluminação entre duas regiões. Na figura 9.11, é possível perceber um exemplo deste tipo de fenômeno, onde dois pontos foram escolhidos na imagem e sua cor foi demonstrada nos pontos “a” e “b”. Os pontos “c” e “d” mostram as cores com os atributos de intensidade e saturação no máximo para que seja possível visualizar melhor a diferença de tonalidade (matiz), que é de 22° . É importante salientar que a imagem do exemplo utiliza uma câmera de última geração, que captura vídeo de alta resolução em formato digital, este tipo de efeito é consideravelmente mais perceptível em câmeras mais antigas e com menos recursos.

Este tipo de problema na captação das imagens resulta muitas vezes na classificação errada da região dominante, pois as regiões, (saturada e não saturada), podem não ser consideradas como parte da mesma região de campo.



Figura 9.11 – Diferença de tonalidade entre dois pontos na imagem, (a) e (b) são os pontos originais e (c) e (d) seus correspondentes de cor pura (saturação e valor em 100%).

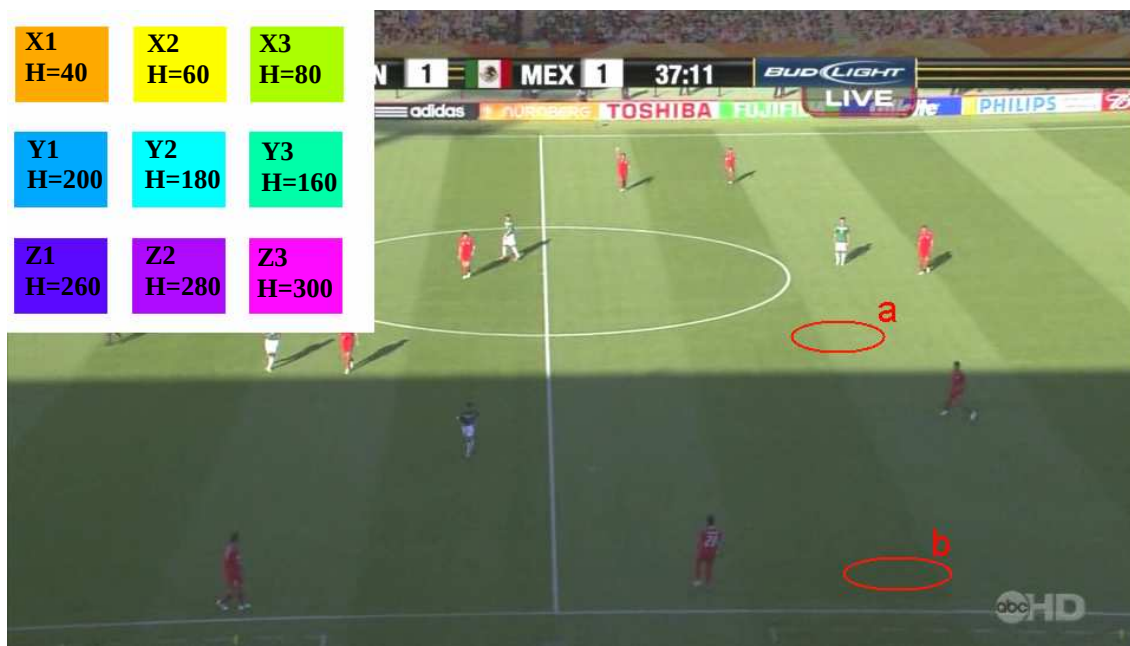


Figura 9.12 – Diferença entre região de campo com e sem sombra.

A figura 9.12 também ilustra o problema citado. A região circulada “a” tem valores de H entre 77° e 81° , e na região “b” entre 102° e 127° , ou seja, uma diferença de mais de 20° . Para ilustrar o que representa a distância de 20° no espaço de cor HSV, no canto superior esquerdo da figura 9.12, existem três (3) linhas de quadrados de valores diferentes de H, (porém de mesmo valor de S e V), e a distância em graus de um quadrado para o outro nas linhas é de apenas 20° . Caso exista diferença de S e V, a impressão visual de distância entre as cores é ainda maior.

Outro tipo de problema é a mistura de cores nas regiões de fronteira entre dois grupos de cores diferentes. A figura 9.14 mostra a ampliação da fração de uma imagem contendo um jogador, e é possível perceber claramente os efeitos deste espalhamento, em especial nos pontos azulados no gramado herdados do calção do jogador. Este efeito é causado por vários fatores, entre eles, compressão, geração de má qualidade, ruídos em alguma etapa da transmissão analógica do sinal, etc.

Em um dos jogos houveram problemas relacionados com a integridade dos dados do vídeo. Não é possível saber se ocorreu na transmissão, na compressão ou na transferência pela internet. O mais provável é ter ocorrido na transmissão uma vez que é um problema comum nas transmissões digitais especialmente em sistemas via satélite. Este tipo de problema é indicado por círculos vermelhos na figura 9.13.

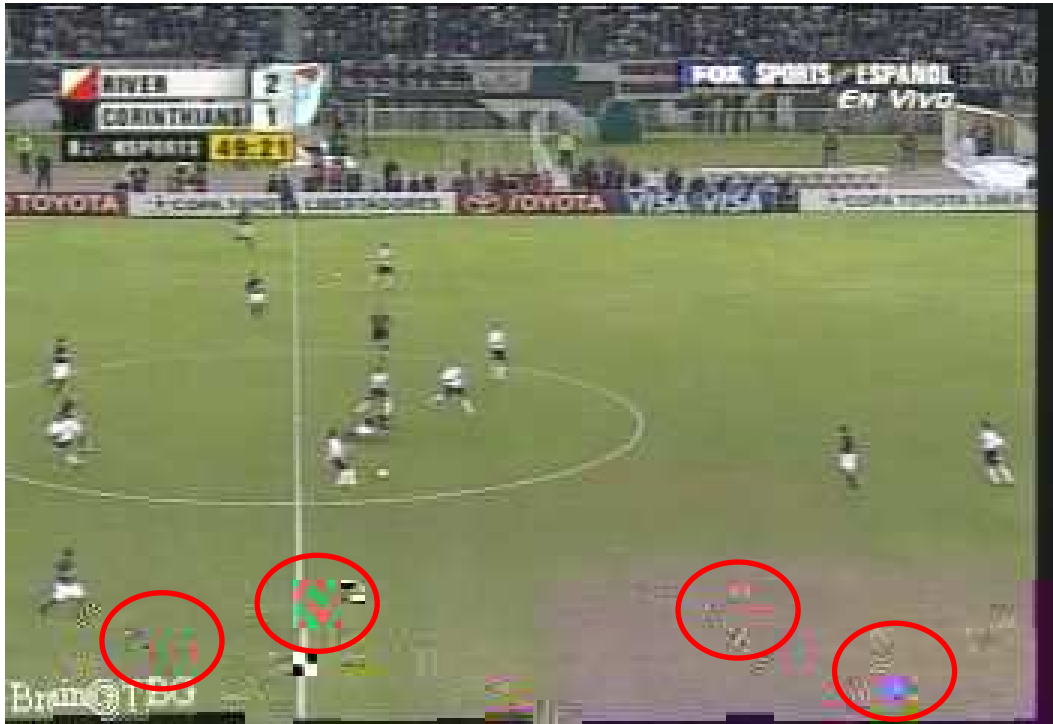


Figura 9.13 – Detalhes do espalhamento das cores



Figura 9.14 – Detalhes do espalhamento das cores

Ao contrário do que pode parecer a princípio, a separação dos jogadores em dois grupos não é uma tarefa elementar, pois apresenta alguns problemas:

- Variação muito grande na tonalidade das cores das camisas (Figura 9.15);
- Dependendo da cena, os números da camisa, a publicidade estampada, ou o calção são muito representativos na região segmentada;

- Quando a região do jogador é muito pequena, a mistura de cores tem uma influência muito grande do campo e toma cores diferentes dos outros jogadores;
- Imagem borrada ou com arrasto (Figura 9.16), normalmente causada pelo movimento dos jogadores;



Figura 9.15 – Diferença das tonalidades de cor de dois jogadores extraídos de uma mesma imagem



Figura 9.16 – Problema de imagem borrada: (a) e (b) Imagens captadas de um jogo, (c) e (d) Zoom dos jogadores marcados em (a) e (b).

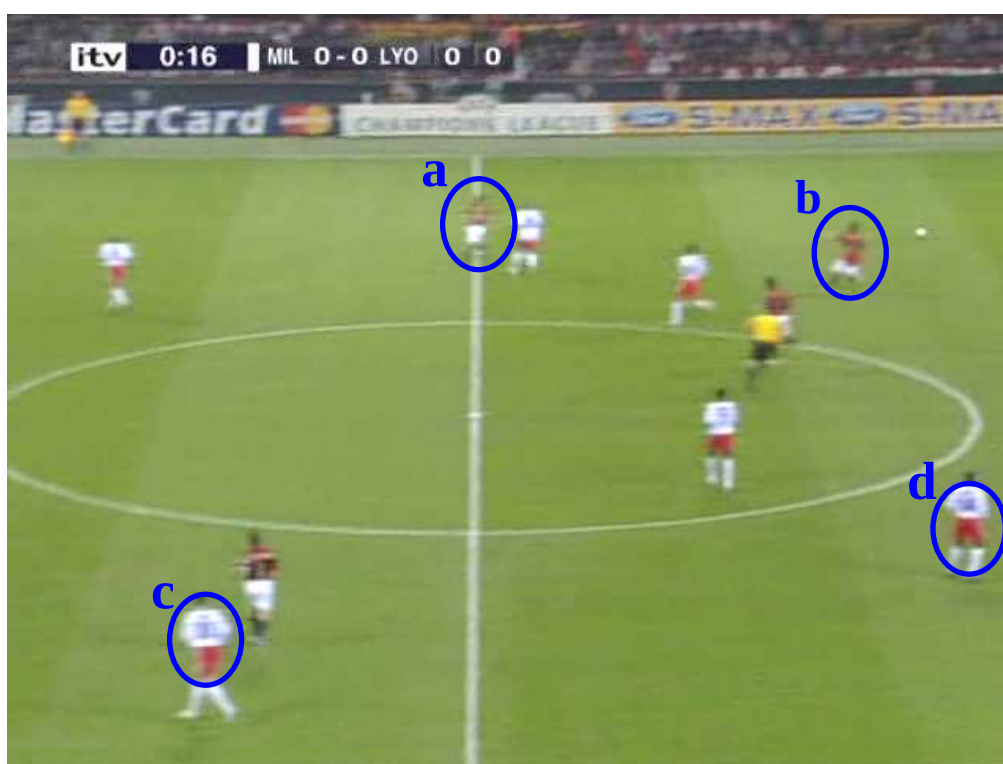


Figura 9.17 – Confusões causadas por ângulo de visão

Outro problema sério encontrado nos testes, diz respeito às cores predominantes dos jogadores conforme o ângulo de visão da câmera. Como pode ser observado na figura 9.17, o jogador (a) tem predominância da cor branca enquanto o (b) do mesmo time tem predominância vermelha. Na mesma cena o mesmo problema ocorre no outro time, onde o jogador (c) tem predominância da cor branca e o d da cor vermelha. Em um caso como este, é bem possível que os jogadores (a) e (c) sejam classificados como pertencentes a um time, e os jogadores (b) e (d) pertencentes a outro. Normalmente no futebol as camisas e calções são escolhidos de maneira que sejam facilmente percebidas as diferenças entre os jogadores de cada time, porém em alguns casos a diferença não é suficiente dado às restrições técnicas diversas. É importante lembrar que mesmo aparentemente diferente, neste caso as tonalidades de vermelho se diferenciam muito pouco quando vistas sob o espaço de cor HSV, na componente H, pois a maior diferença entre os dois vermelhos é dada pela variação do S e do V. Neste caso especificamente, tanto os calções vermelhos de um time quanto as camisas vermelhas do outro são compostos por valores de H de 350° a 30° .

Uma dificuldade encontrada, e que também refletiu nos resultados da classificação, é quando a camisa de um time tem mais de uma cor, e as tomadas são feitas de ângulos diferentes do jogador. Na figura 9.18 podem ser vistos dois jogadores de mesmo time classificados em grupos diferentes, pois a visão da camisa por trás é muito diferente da visão pelo lado.

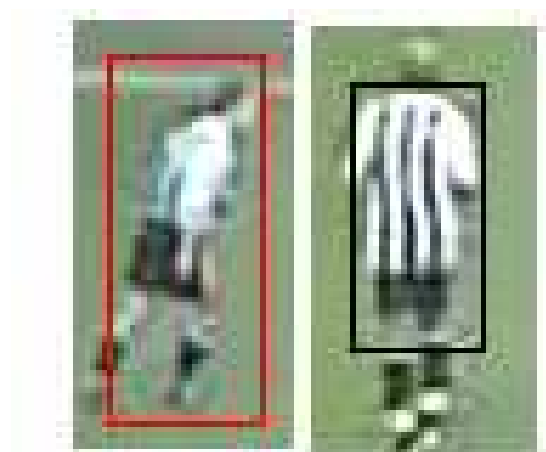


Figura 9.18 – Diferença da camisa dos jogadores em um mesmo quadro

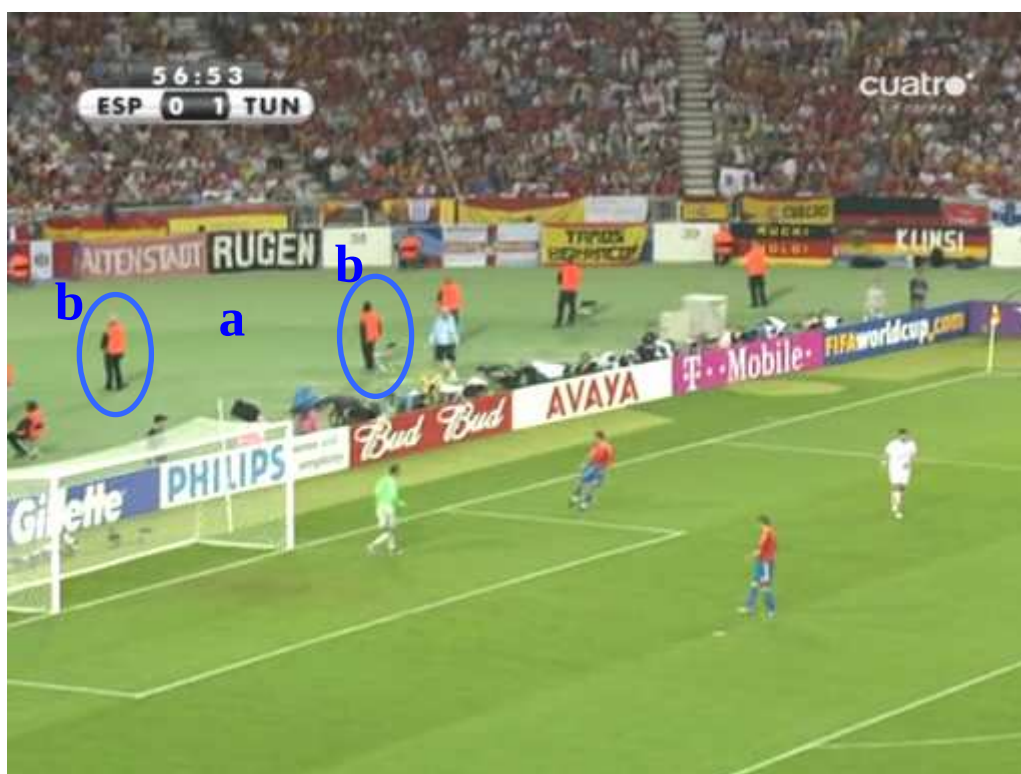


Figura 9.19 – Região externa (a) classificada como campo, e seguranças classificados como jogadores (b)

Um problema curioso que também causou perda no desempenho do sistema, foi em casos como o da figura 9.19, onde a parte posterior do campo é gramada (a), logo acaba sendo segmentada como campo, e os seguranças vestindo alaranjado que lá estão (b), acabam por ser segmentados como jogadores. Mesmo sendo uma prova do correto funcionamento do algoritmo, que separou bem a região predominante, e as regiões nela inseridas, este tipo de imagem teve as partes externas consideradas regiões adjacentes independente da cor.

Capítulo 10

Conclusão e Trabalhos Futuros

Este trabalho apresentou uma nova abordagem automática para o processo de segmentação e classificação em vídeos esportivos. A escolha do futebol como o esporte a ser testado se mostrou acertada, pois possibilitou obter um conjunto variado de vídeos, proporcionando muitas situações diferentes e inesperadas, mencionadas no capítulo de experimentos. Esta variedade também possibilitou a obtenção de resultados e análises que não seriam possíveis com o uso de um conjunto de vídeos similares.

O algoritmo desenvolvido recebe como entrada o quadro de um vídeo, a saída é um quadro com os jogadores e demais elementos em campo marcados e classificados por grupos semânticos baseados em cor, ou seja, jogadores de um time de uma cor, e jogadores de outro time de outra, juiz, goleiros, e demais objetos em campo com cores distintas.

O sistema não emprega uma base de conhecimento, e é orientado apenas pela premissa de que o campo é a região predominante num vídeo esportivo, e que os jogadores e o juiz estão neste campo.

O índice de qualidade utilizado para a análise dos resultados foi gerado tanto priorizando determinadas etapas quanto dando o mesmo peso para todas, e em todos os casos os resultados foram realmente promissores, fato que pode ser comprovado pelos dados mostrados no apêndice C.

Apesar da grande diferença de qualidade entre os vídeos utilizados, e da variedade de problemas encontrados, o resultado final ficou acima das expectativas. Um bom exemplo disso foi o jogo 1, onde foi atingido índice de 83,70%, lembrando que este jogo tem imagens de péssima qualidade, em uma resolução de 352x288 pixels. Outro jogo de má qualidade que teve resultado parecido foi o de número 9, que apesar

de melhor resolução também tem qualidade de imagem ruim, e atingiu os mesmos 83,70%. Um jogo de boa qualidade, porém de resolução ainda baixa (640x480), como o jogo 7, obteve índice 94,55%, e outro de alta resolução (960x544), como o jogo 13, obteve índice de 91,56%. Os vídeos com os piores resultados, obtiveram índices de 75,92% e 80,96%, resultados que para grande parte das aplicações ainda são aceitáveis.

O ponto forte do trabalho foi sem dúvida a segmentação automática de região predominante na imagem que trouxe resultados empolgantes, tanto nos vídeos de má quanto nos de alta qualidade. Isto é demonstrado pelo percentual do índice de acerto dos pixels de campo, que teve 12 dos 14 vídeos com índice acima de 95%, sendo 8 deles acima de 98%. Os piores índices de acerto de pixels de campo foram de 63,48% e 78,68%, em dois jogos onde sombras no campo causavam um contraste muito grande nas cores da grama (entre as regiões com e sem a incidência do sol). Porém, nem todos os quadros que tinham estas sombras tiveram falsas detecções de campo, melhorias na abordagem de análise dos histogramas de acumulação deverão resolver estes problemas.

A diversidade dos vídeos utilizados e seus resultados consistentes demonstraram a robustez e a eficiência do método. Os desafios e problemas que surgiram durante os experimentos, serviram para motivar a busca por melhores soluções, e para estimular a continuidade deste desenvolvimento.

A criação de um mecanismo realmente automático, capaz de processar uma variedade de tipos e qualidades de imagem, sem a aplicação de parâmetros, heurísticas, ou intervenção de um operador, é outro ponto de destaque. Isto representa uma sólida contribuição científica, uma vez que os conceitos aqui definidos poderão ser utilizados em uma ampla variedade de trabalhos.

Trabalhos futuros podem melhorar da extração de informação dos elementos segmentados para a classificação dos grupos. A simples extração dos pixels como foi utilizada perde desempenho quando submetida a imagens de má qualidade, onde a mistura das cores dificulta a separação em grupos realmente distintos entre si. É possível que uma filtragem, ou uma identificação mais precisa da posição do jogador tornem o processo de classificação muito mais preciso. É possível também melhorar a detecção automática do campo, corrigindo discrepâncias em campos com forte incidência de sol e com sombra de cores muito diferentes, e também resolver o problema da fusão de jogadores muito próximos.

Referências Bibliográficas

- [ANDRADE03] ANDRADE, Ernesto L.; KHAN, Ekram, WOODS, John C.; GHANBARI, Mohammed. *Segmentation and Tracking For Interactive Sport Scenes Using Region Adjacency Graphs, Picture Trees and Prior Information*. IEE IEE Visual Information Engineering (VIE) 2003, July, 2003, University of Surrey, Guildford, UK p. 45-48
- [ANDRADE04] ANDRADE , Lucio P. de. *Procedimento Iterativo de Agrupamento de Dados*. Tese apresentada ao Programa de Pós-Graduação de Engenharia da Universidade Federal do Rio de Janeiro, UFRJ, Rio de Janeiro, 2004
- [BUSIN04] BUSIN, L.; VANDENBROUCKE, N.; MACAIRE, L.; POSTAIRE, J-G. *Color space selection for unsupervised color image segmentation by histogram multithresholding*, IEEE 2004
- [CABROL05] CABROL, Aymeric de; BONNIN, Patrick; COSTIS, Thomas; HUGEL, Vincent; BLAZEVIC Pierre; BOUCHEFRA, Kamel Bouchefra. *A New Video Rate Region Color Segmentation and Classification for Sony Legged RoboCup Application*. In RoboCup 2005: Robot Soccer World Cup IX, 2005
- [CSMSU06] Graphics & Media Lab, The Laboratory of computer graphics at dep. Of CS MSU. <http://graphics.cs.msu.su>, acessado em 14 de Julho de 2006.
- [CUMANI91] CUMANI, A. *A note on the gradient of multi-image edge detection in multispectral images*. Computer Vision Graphics Image Process 53, 1991, p.40–51
- [CHANG02] CHANG, Jeng-Horng; FAN, Kuo-Chin; CHANG, Yang-Lang. *Multi-modal gray-level histogram modeling and decomposition*. Image and Vision Computing 20, 2002, p. 203-216

- [CHEN02] CHEN, Tie Qi; LU, Yi. *Color image segmentation - an inovative approach*. Patern Recognition 35, 2002, p. 395-405
- [CHENG03] CHENG, H. D.; LI, J. *Fuzzy homogeneity and scale-space approach to color image segmentation*. Patern Recognition 36, 2003, p. 1545-1562
- [DREW00] DREW, Mark S.; LI, Ze-Nian; ZHONG Xiang. *Video dissolve and wipe detection via spatio-temporal images of chromatic histogram differences*. International Conference on Image Processing (ICIP 2000), vol. 3, p. 929–932
- [DUARTE02] DUARTE, Glaucius Décio. *Recuperação de Imagens Digitais por Similaridade Cromática*. Submetido ao VII Simpósio de Informática e II Mostra Regional de Software Acadêmico, PUCRS, Câmpus Uruguaiana, novembro de 2002 (não publicado).
- [EKIN02] EKIN, Ahmet; TEKALP, A. Murat. *Robust Dominant Color Region Detection with Applications to Sports Video Analysis*. Department of Electrical and Computer Engineering, University of Rochester, and College of Engineering, Koc University, Istanbul, Turkey 2002
- [FACON96] FACON, Jaques. *Morfologia Matemática: Teoria e Exemplos*. Editora Universitária Champagnat da Pontifícia Universidade Católica do Paraná, Curitiba, PR, 1996
- [FIGUEROA04] FIGUEROA, Pascual; LEITE, Neucimar; BARROS, Ricardo M. L.; COHEN, Isaac; MEDIONI, Gerard. *Tracking soccer players using the graph representation*. Proceedings of the Pattern Recognition, 17th International Conference on (ICPR'04) Volume 4, 2004, p. 787-790
- [GABRIEL03] GABRIEL, Pierre F.; VERLY, Jacques G.; PIATER, Justus H.GENON, André. *The State of the Art in Multiple Object Tracking Under Occlusion in Video Sequences*. Acivs 2003 Advanced Concepts for Intelligent Vision Systems, Ghent, Belgium, 2003

- [GU04] GU, Liexian; DING, Xiaoqing; HUA, Xian-ShengOnline. *Play Segmentation for Broadcasted American Football TV Programs*. Department of Electronic Engineering, Tsinghua University, Beijing, China / Microsoft Research Asia, Beijing, China, 2004
- [GUIMARÃES01] GUIMARÃES, Silvio Jamil Ferzoli; COUPRIE, Michel; LEITE, Neucimar Jerônimo; ARAÚJO, Arnaldo de Albuquerque. *Análise do vídeo a partir da sua simplificação*. Anais do II Workshop em Tratamento de Imagens, NPDI/DCC/UFGM, 2001
- [HAN01] J. HAN, M. KAMBER. *Data Mining: Concepts and Techniques*, Morgan Kaufmann. San Francisco, 2001 p. 346–389.
- [ISMAIL89] ISMAIL, M.A.,KAMEL, M.S. *Multidimensional data clustering utilizing hybrid search strategies*. Pattern Recognition 22, 1989, p. 75–89
- [KOPRINSKA00] KOPRINSKA, Irena; CARRATO, Sergio. *Temporal video segmentation: A survey*. Signal Processing: Image Communication 16, 2001, p. 477-500, 2000
- [KOZAMERNIK04] KOZAMERNIK, F.; STEINMAN, V.; SUNNA, P.; WYCKENS, E. *SAMVIQ – A New EBU Methodology for Video Quality Evaluations in Multimedia*. IBC 2004, Amsterdam, p. 191-202.
- [KURUGOLLU01] KURUGOLLU, F.; SANKUR, B; HARMANCI, A. E. *Color segmentation using histogram multithresholding and fusion*. Image and Vision Computing 19, 2001, p. 915-928
- [LEFÈVRE02] LEFÈVRE, Sébastien; MERCIER, Loïc; TIBERGHIEN, Vincent; VINCENT, Nicole. *Multiresolution Color Image Segmentation Applied to Background Extraction in Outdoor Images*. RFAI Publication, IS&T European Conference on Color in Graphics, Image and Vision, Poitiers (France), April 2002, p. 363-367

- [LIAO05] LIAO, T. Warren. *Clustering of time series data - a survey*. Pattern Recognition 38(11), 2005, p. 1857-1874
- [LOWE04] LOWE, David G. *Distinctive image features from scale-invariant keypoints*. University of British Columbia, 2004
- [MACQUEEN67] MACQUEEN, J. B. *Some Methods for classification and Analysis of Multivariate Observations*. Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, University of California Press, 1967, p. 281-297
- [MEHROTRA03] EKIN, Ahmet; TEKALP, A. Murat; MEHROTRA Rajiv. *Automatic Soccer Video Analysis and Summarization*. Image Processing, IEEE Transactions on, Volume: 12, Issue: 7, 2003, p. 796-807
- [MITRA98] Lucchese, L.; MITRA, S. K. *An algorithm for fast segmentation os color images*. University of Padua, 1998
- [NAKAMURA99] NAKAMURA, Takayuki; OGASAWARA, Tsukasa. *On-Line Visual Learning Method for Color Image Segmentation and Object Tracking*. Intelligent Robots and Systems, 1999. IROS '99. p. 222-228 vol.1
- [NEEDHAM01] NEEDHAM, Chris J.; BOYLE, Roger D. *Tracking multiple sports players through occlusion, congestion and scale*. Proceedings British Machine Vision Conference, Manchester, UK, 2001, vol. 1, p.93-102
- [OHTA80] OHTA, Y.I., KANADE, T., SAKAI, T. *Color information for region segmentation*. Computer Graphics Image Process 13, 1980, p.222–241.
- [OZYILDIZ02] OZYILDIZ, Ercan; KRAHNSTOEVER, Nils; SHARMA Rajeev. *Adaptive Texture and Color Segmentation for Tracking Moving Objects*. Pattern Recognition 35, 2002 p. 2013-2029

- [PAINTNET06] PAINT .NET, <http://www.eecs.wsu.edu/paint.net/pdnabout.html>, acessado em 14 de Julho de 2006.
- [PANJAWANI95] PANJAWANI, D.K., HEALEY, G. *Markov random field models for unsupervised segmentation of textured color images*. IEEE Transactions on Pattern Analysis and Machine Intelligence 17, 1995 p.939–954.
- [PAPAMARKOS00] PAPAMARKOS, N.;STROUTHOPOLOS, C.; ANDREADIS, I. *Multithresholding of color and gray-level images through a neural network technique*. Image and Vision Computing 19, 2000, p. 213-222
- [RODRIGUES03] RODRIGUES, Paulo Sérgio Silva; ARAUJO, Arnaldo de Albuquerque. *Um Modelo Bayesiano Combinando Análise Semântica Latente e Atributos Espaciais para Recuperação de Informação Visual*. Tese apresentada ao Curso de Pós-graduação em Ciência da Computação da Universidade Federal de Minas Gerais, UFMG, Belo Horizonte, MG, 2003
- [SANTOS04] SANTOS, Thiago Teixeira. *Segmentação automática de tomadas em vídeo*. Dissertação apresentada ao Instituto de Matemática e Estatística da USP, 2004
- [SHAFARENKO98] SHAFARENKO, L.,PETROU, M., KITTLER, J.V. *Histogram based segmentation in a perceptually uniform color space*. IEEE Transactions on Image Process 7, 1998 p.1354–1358.
- [SOILLE99] SOILLE, P., “*Morphological Image Analysis - Principle and Applications*”. Springer-Verlag Berlin and Heidelberg GmbH & Co. K, 1999.
- [SOUTO00] Souto, R.P. *Segmentação de imagem multispectral utilizando-se do atributo matiz*. Instituto Nacional de Pesquisas Espaciais, 2000
- [SOUTO00] SOUTO, Roberto Pinto. *Segmentação de imagem multispectral utilizando-se do atributo matiz*. São José dos Campos: INPE (Instituto Nacional de Pesquisas Espaciais), 2000

- [SUN00] Cheng, H.D.; SUN, H. *A hierarchical approach to color image segmentation using homogeneity*. IEEE Transactions on image processing, 2000, p. 2071-2082
- [TAJKAHASHI95] K., NAKATANI, H., ABE, K. *Color image segmentation using ISODATA clustering method*. Second Asian Conference on Computer Vision, Singapore, vol. 1, 1995, p.523–527
- [THIRDE03] THIRDE, D.J.; JONES G.A.; FLACK, J. *Spatio-Temporal Semantic Object Segmentation using Probabilistic Sub-Object Regions*. British Machine Vision Conference, 2003, p. 163-172
- [TRÉMEAU00] TRÉMEAU, A., COLANTONI, P. *Regions adjacency graph applied to color image segmentation*. IEEE Transactions on Image Process 9, 2000 p.735–744.
- [VANDENBROUCKE03] VANDENBROUCKE, Nicolas; MACAIRE, Ludovic; POSTAIRE, Jack-Gérard. *Color image segmentation by pixel classification in an adapted hybrid color space. Application to soccer image analysis*. Computer Vision and Image Understanding 90, 2003, p. 190–216
- [VIRTUALDUB06] VIRTUAL DUB, <http://www.virtualdub.org>, acessado em 14 de Julho de 2006.
- [WANG01] WANG, James; LI, Jia; GRAY, Robert M.; Wiederhold, Gio. *Unsupervised Multiresolution Segmentation for Images with Low Depth of Field*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001
- [WANG02] WANG, Zhou; BOVIK, Alan C., LU, Ligang. *Why is Image Quality Assessment So Difficult?* IEEE International Conference on Acoustics, Speech, & Signal Processing, 2002
- [WANG03] WANG, Hanzi; SUTER, David. *Color Image Segmentation Using Global Information and Local Homogeneity*. 7th International Conference on

Digital Image Computing: Techniques and Applications (DICTA'03), Sydney 2003, p. 89-98

[WANG97] W. WANG, J. Yang, R. Muntz, R. *STING: a statistical information grid approach to spatial data mining*. Proceedings of the 1997 International Conference on Very Large Data Base (VLDB'97), Athens, Greek, 1997, pp. 186–195.

[WIKIPEDIA06] Wikipedia A Enciclopédia Livre. <http://www.wikipedia.org>, acessado em 14 de Julho de 2006.

[WU04] WU, B.F.; CHEN, Y.L.; CHIU, C.C. *Efficient implementation of several multilevel thresholding algorithms using a combinatorial scheme*. Department of Electrical and Control Engineering, National Chiao Tung University, 2004

[YEN95] YEN, J.C.; CHANG, F.J; CHANG, S. *A new criterion for automatic multilevel thresholding*. IEEE Transactions on image processing, 1995 vol 4, no 3

[ZENZO86] ZENZO, S.D. *A note on the gradient of a multi-image*, Comput. Vision Graphics Image Process. 33, 1986, p.116–125

Apêndice A

Planilhas de Resultados



Figura A.1– Exemplo de Classificação

Vídeo: 2006-01-02 Newcastle vs middlesbrough

Resolução: 640x480

Qualidade: BOM

CIM	CIA	RAIM	RAIA	TRS	RCC	TJ	JCA	JCC	JNS	
291906	297161	0	0	10	10	6	1	5	0	
225334	230003	66263	68594	13	13	9	2	7	0	
165388	169491	131957	132321	7	7	9	4	5	0	
209191	212823	89550	90184	7	5	6	1	3	0	
294561	296838	0	0	10	8	6	2	2	0	
292429	295606	0	0	9	6	7	0	4	0	
254773	257569	44293	44800	9	5	6	1	3	0	
189258	193188	109925	110578	5	3	5	3	2	0	
270209	277894	24252	24793	4	4	7	3	4	0	
232066	238318	64506	64998	9	6	5	0	3	0	
191365	200837	103115	103404	7	4	7	1	4	0	
286665	290030	702	750	7	5	6	1	3	0	
199061	205429	97336	97725	19	19	6	2	4	0	
204999	208563	93693	94056	11	7	10	2	4	0	
266502	271955	28910	29373	15	11	11	1	7	0	
232810	234999	69410	70503	6	4	4	2	1	0	
247390	250722	52479	54703	8	7	4	2	1	0	
285945	292632	0	0	18	15	14	2	10	0	
185142	189643	112044	112324	8	7	8	1	6	0	
235437	239032	62067	63806	4	4	4	0	4	0	
4760431	4852733	1150502	1162912	186	150	140	31	82	0	Total

Tabela A.1 – Tabela de dados

Vídeo: 2006-01-02 Newcastle vs middlesbrough
 Resolução: 640x480
 Qualidade: BOM

CMPO	ADJ	JSC1	JSC2	JCC1	JCC2	RCCP	IQ1	IQ2	
98,23%	100,00%	83,33%	100,00%	83,33%	100,00%	100,00%	96,05%	99,38%	
97,97%	96,60%	77,78%	100,00%	77,78%	100,00%	100,00%	93,66%	98,10%	
97,58%	99,72%	55,56%	100,00%	55,56%	100,00%	100,00%	90,17%	99,06%	
98,29%	99,30%	83,33%	100,00%	50,00%	60,00%	71,43%	89,63%	92,30%	
99,23%	100,00%	66,67%	100,00%	33,33%	50,00%	80,00%	87,73%	92,73%	
98,93%	100,00%	100,00%	100,00%	57,14%	57,14%	66,67%	92,00%	92,00%	
98,91%	98,87%	83,33%	100,00%	50,00%	60,00%	55,56%	88,11%	90,78%	
97,97%	99,41%	40,00%	100,00%	40,00%	100,00%	60,00%	83,08%	95,08%	
97,23%	97,82%	57,14%	100,00%	57,14%	100,00%	100,00%	89,70%	98,27%	
97,38%	99,24%	100,00%	100,00%	60,00%	60,00%	66,67%	91,48%	91,48%	
95,28%	99,72%	85,71%	100,00%	57,14%	66,67%	57,14%	88,25%	90,63%	
98,84%	93,60%	83,33%	100,00%	50,00%	60,00%	71,43%	87,83%	90,50%	
96,90%	99,60%	66,67%	100,00%	66,67%	100,00%	100,00%	92,11%	98,78%	
98,29%	99,61%	80,00%	100,00%	40,00%	50,00%	63,64%	87,63%	90,63%	
97,99%	98,42%	90,91%	100,00%	63,64%	70,00%	73,33%	91,53%	93,08%	
99,07%	98,45%	50,00%	100,00%	25,00%	50,00%	66,67%	83,30%	90,80%	
98,67%	95,93%	50,00%	100,00%	25,00%	50,00%	87,50%	84,36%	91,86%	
97,71%	100,00%	85,71%	100,00%	71,43%	83,33%	83,33%	93,25%	95,87%	
97,63%	99,75%	87,50%	100,00%	75,00%	85,71%	87,50%	94,08%	96,40%	
98,50%	97,27%	100,00%	100,00%	100,00%	100,00%	100,00%	98,52%	98,52%	
98,10%	98,93%	77,86%	100,00%	58,57%	75,23%	80,65%	90,67%	94,55%	TOTAL

Tabela A.2 – Tabela de Resultados



Figura A.2 – Exemplo de Classificação

Vídeo: Confeds Cup Final Brazil vs Argentina 1st Half.avi

Resolução: 352x288

Qualidade: RUIM

Qualidade: RAIM									
CIM	CIA	RAIM	RAIA	TRS	RCC	TJ	JCA	JCC	JNS
74725	76202	20690	23480	13	11	14	1	9	2
66670	67980	29431	32560	5	3	8	3	3	0
83401	84107	13661	16286	16	10	10	1	3	2
85862	86053	6299	12597	10	6	15	0	4	10
49084	49507	35576	50363	18	15	11	1	7	0
74781	76589	20708	23204	11	6	10	1	5	0
70188	71472	26456	27967	11	9	17	1	9	5
86180	87937	10725	12752	9	4	9	0	4	0
96738	98938	0	0	9	4	9	0	4	0
80072	82218	12104	17124	17	8	14	0	6	1
80979	82179	9047	16242	10	5	15	1	4	8
94041	94339	1988	3468	18	15	16	0	13	1
61521	63685	26818	35218	21	15	11	2	6	0
71212	73250	24556	26676	8	4	10	2	4	0
65814	66958	29564	33820	5	3	5	1	3	1
67655	69116	28015	30982	12	6	14	0	6	0
97869	100285	0	0	4	2	2	0	2	0
66682	66796	22558	33851	9	6	5	0	3	0
75012	76367	22003	24372	3	2	6	0	2	3
79002	79540	15770	19866	8	5	7	0	5	1
1527488	1553518	355969	440828	217	139	208	14	102	34
Total									

Tabela A.3 – Tabela de dados

Vídeo: Confeds Cup Final Brazil vs Argentina 1st Half.avi
 Resolução: 352x288
 Qualidade: RUIM

CMPO	ADJ	JSC1	JSC2	JCC1	JCC2	RCCP	IQ1	IQ2	
98,06%	88,12%	78,57%	84,62%	75,00%	81,82%	84,62%	88,98%	90,27%	
98,07%	90,39%	62,50%	100,00%	37,50%	60,00%	60,00%	81,96%	87,96%	
99,16%	83,88%	70,00%	77,78%	37,50%	42,86%	62,50%	81,06%	82,38%	
99,78%	50,00%	33,33%	33,33%	80,00%	80,00%	60,00%	69,76%	69,76%	
99,15%	70,64%	90,91%	100,00%	63,64%	70,00%	83,33%	83,21%	84,76%	
97,64%	89,24%	90,00%	100,00%	50,00%	55,56%	54,55%	84,86%	86,42%	
98,20%	94,60%	64,71%	68,75%	75,00%	81,82%	81,82%	89,63%	90,72%	
98,00%	84,10%	100,00%	100,00%	44,44%	44,44%	44,44%	82,63%	82,63%	
97,78%	100,00%	100,00%	100,00%	44,44%	44,44%	44,44%	88,11%	88,11%	
97,39%	70,68%	92,86%	92,86%	46,15%	46,15%	47,06%	77,43%	77,43%	
98,54%	55,70%	40,00%	42,86%	57,14%	66,67%	50,00%	68,70%	69,94%	
99,68%	57,32%	93,75%	93,75%	86,67%	86,67%	83,33%	81,33%	81,33%	
96,60%	76,15%	81,82%	100,00%	54,55%	66,67%	71,43%	81,24%	84,27%	
97,22%	92,05%	80,00%	100,00%	40,00%	50,00%	50,00%	83,24%	86,24%	
98,29%	87,42%	60,00%	75,00%	75,00%	100,00%	60,00%	84,50%	88,50%	
97,89%	90,42%	100,00%	100,00%	42,86%	42,86%	50,00%	85,19%	85,19%	
97,59%	100,00%	100,00%	100,00%	100,00%	100,00%	50,00%	94,16%	94,16%	
99,83%	66,64%	100,00%	100,00%	60,00%	60,00%	66,67%	80,93%	80,93%	
98,23%	90,28%	50,00%	50,00%	66,67%	66,67%	66,67%	84,31%	84,31%	
99,32%	79,38%	85,71%	85,71%	83,33%	83,33%	62,50%	85,70%	85,70%	
98,32%	80,75%	76,92%	82,47%	58,62%	63,75%	64,06%	82,64%	83,70%	TOTAL

Tabela A.4 – Tabela de Resultados



Figura A.3 – Exemplo de Classificação

Vídeo: FIFA.World.Cup.2006.Espanya.Vs.Tunez.[www.tensiontorrent.com].CD2.avi

Resolução: 512x384

Qualidade: BOM

CIM	CIA	RAIM	RAIA	TRS	RCC	TJ	JCA	JCC	JNS	
97975	99339	78472	94251	22	15	13	1	5	0	
131514	132564	40646	61477	13	7	12	2	5	0	
161237	162947	21134	30838	16	10	12	0	6	0	
165160	166286	18882	27931	20	17	12	0	8	1	
171885	173326	9058	21490	16	12	8	0	4	0	
119607	120638	55646	75158	12	8	10	1	5	0	
146403	148017	32257	46267	26	16	11	0	6	0	
143061	145344	34285	48731	24	17	9	1	5	0	
149780	150602	35530	44399	16	14	8	1	5	0	
85090	86394	65527	109328	16	8	4	1	1	0	
145735	146419	41040	48640	13	10	6	0	4	0	
138387	140272	36126	53760	25	18	10	0	3	0	
132588	135674	44686	57787	20	14	9	1	2	0	
141326	143998	35389	50861	29	26	7	2	2	0	
149856	151267	26561	43008	21	16	12	1	8	0	
143517	144809	30857	48808	18	10	11	1	5	0	
157311	159526	5665	33249	20	14	12	2	3	1	
164962	166355	20842	28359	13	8	7	0	3	0	
140890	144169	32471	48862	23	18	9	1	3	0	
127261	127870	55892	66560	18	13	11	1	6	0	
2813545	2845816	720966	1039764	381	271	193	16	89	2	Total

Tabela A.5 – Tabela de dados

Vídeo: FIFA.World.Cup.2006.Espanya.Vs.Tunez.[www.tensiontorrent.com].CD2.avi
 Resolução: 512x384
 Qualidade: BOM

CMPO	ADJ	JSC1	JSC2	JCC1	JCC2	RCCP	IQ1	IQ2	
98,63%	83,26%	92,31%	100,00%	38,46%	41,67%	68,18%	83,56%	84,64%	
99,21%	66,12%	83,33%	100,00%	41,67%	50,00%	53,85%	75,75%	78,25%	
98,95%	68,53%	100,00%	100,00%	50,00%	50,00%	62,50%	79,87%	79,87%	
99,32%	67,60%	91,67%	91,67%	72,73%	72,73%	85,00%	83,36%	83,36%	
99,17%	42,15%	100,00%	100,00%	50,00%	50,00%	75,00%	71,96%	71,96%	
99,15%	74,04%	90,00%	100,00%	50,00%	55,56%	66,67%	81,28%	82,84%	
98,91%	69,72%	100,00%	100,00%	54,55%	54,55%	61,54%	80,63%	80,63%	
98,43%	70,36%	88,89%	100,00%	55,56%	62,50%	70,83%	80,60%	82,41%	
99,45%	80,02%	87,50%	100,00%	62,50%	71,43%	87,50%	86,57%	88,71%	
98,49%	59,94%	75,00%	100,00%	25,00%	33,33%	50,00%	70,45%	73,78%	
99,53%	84,38%	100,00%	100,00%	66,67%	66,67%	76,92%	88,73%	88,73%	
98,66%	67,20%	100,00%	100,00%	30,00%	30,00%	72,00%	78,25%	78,25%	
97,73%	77,33%	88,89%	100,00%	22,22%	25,00%	70,00%	79,38%	80,77%	
98,14%	69,58%	71,43%	100,00%	28,57%	40,00%	89,66%	77,67%	81,67%	
99,07%	61,76%	91,67%	100,00%	66,67%	72,73%	76,19%	79,74%	81,18%	
99,11%	63,22%	90,91%	100,00%	45,45%	50,00%	55,56%	76,01%	77,37%	
98,61%	17,04%	75,00%	90,00%	27,27%	33,33%	70,00%	57,70%	59,81%	
99,16%	73,49%	100,00%	100,00%	42,86%	42,86%	61,54%	80,87%	80,87%	
97,73%	66,45%	88,89%	100,00%	33,33%	37,50%	78,26%	77,51%	79,04%	
99,52%	83,97%	90,91%	100,00%	54,55%	60,00%	72,22%	85,99%	87,45%	
98,87%	69,34%	90,67%	98,87%	46,60%	50,86%	71,13%	79,71%	80,96%	TOTAL

Tabela A.6 – Tabela de Resultados



Figura A.4 – Exemplo de Classificação

Vídeo: Final.Copa.del.Rey.2006.Espanyol-Zaragoza.HQTVRip.avi

Resolução: 512x380

Qualidade: BOM

CIM	CIA	RAIM	RAIA	TRS	RCC	TJ	JCA	JCC	JNS
152958	156264	35683	36251	4	3	3	0	3	0
164479	169259	14446	17220	12	6	11	0	4	1
104721	106540	84280	84929	6	4	10	6	2	0
97382	98985	93891	94061	9	5	10	4	2	0
86761	87775	102556	104089	12	6	13	4	3	0
119621	122203	70093	70226	6	5	6	3	2	0
119051	122551	69758	69785	4	2	3	0	2	0
147454	151491	39224	39844	7	6	7	1	6	0
112139	118391	69731	69902	9	4	7	1	1	0
120317	125533	63462	63530	9	6	10	3	6	0
183247	185216	1962	2560	15	8	10	1	2	0
126461	127555	65286	65429	4	2	5	2	2	0
93569	95911	93089	93786	9	6	12	4	6	0
55741	56574	133581	136213	4	4	7	3	4	0
43266	44620	144841	149159	2	2	4	3	1	0
47341	49067	140283	144638	3	2	4	2	2	0
140226	146033	41870	41930	15	9	13	1	9	0
125926	129546	60372	60416	11	7	10	0	7	0
111819	116797	74191	74221	6	3	5	0	3	0
120174	123703	68361	68380	6	4	5	0	4	0
97581	101580	90905	90960	4	2	3	0	2	0
134114	134901	55855	57612	11	7	12	1	7	1
123008	124560	63488	64021	7	5	8	3	1	1
119612	120713	71218	72148	9	4	9	2	4	0
39671	40440	152107	153530	3	2	2	0	2	0
181497	184019	3265	3345	8	4	8	1	4	0
96473	97192	90960	95167	8	5	7	1	5	0
180546	181143	7228	8455	9	5	8	0	5	0
139393	141311	48751	49656	12	9	12	1	9	0
146375	149919	38611	38670	13	9	11	1	7	0
109378	110441	80577	81465	9	6	9	1	6	0
3640301	3720233	2169925	2201598	246	152	244	49	123	3
									Total

Tabela A.7 – Tabela de dados

Vídeo: Final.Copa.del.Rey.2006.Espanyol-Zaragoza.HQTVRip.avi
 Resolução: 512x380
 Qualidade: BOM

CMPO	ADJ	JSC1	JSC2	JCC1	JCC2	RCCP	IQ1	IQ2	
97,88%	98,43%	100,00%	100,00%	100,00%	100,00%	75,00%	96,21%	96,21%	
97,18%	83,89%	90,91%	90,91%	40,00%	40,00%	50,00%	81,46%	81,46%	
98,29%	99,24%	40,00%	100,00%	20,00%	50,00%	66,67%	81,80%	90,80%	
98,38%	99,82%	60,00%	100,00%	20,00%	33,33%	55,56%	82,93%	88,26%	
98,84%	98,53%	69,23%	100,00%	23,08%	33,33%	50,00%	83,31%	87,41%	
97,89%	99,81%	50,00%	100,00%	33,33%	66,67%	83,33%	85,86%	94,19%	
97,14%	99,96%	100,00%	100,00%	66,67%	66,67%	50,00%	90,65%	90,65%	
97,34%	98,44%	85,71%	100,00%	85,71%	100,00%	85,71%	94,24%	97,09%	
94,72%	99,76%	85,71%	100,00%	14,29%	16,67%	44,44%	82,51%	84,18%	
95,84%	99,89%	70,00%	100,00%	60,00%	85,71%	66,67%	88,17%	93,75%	
98,94%	76,64%	90,00%	100,00%	20,00%	22,22%	53,33%	77,79%	79,01%	
99,14%	99,78%	60,00%	100,00%	40,00%	66,67%	50,00%	84,62%	91,29%	
97,56%	99,26%	66,67%	100,00%	50,00%	75,00%	66,67%	87,22%	93,05%	
98,53%	98,07%	57,14%	100,00%	57,14%	100,00%	100,00%	90,24%	98,81%	
96,97%	97,11%	25,00%	100,00%	25,00%	100,00%	100,00%	82,92%	97,92%	
96,48%	96,99%	50,00%	100,00%	50,00%	100,00%	66,67%	84,38%	94,38%	
96,02%	99,86%	92,31%	100,00%	69,23%	75,00%	60,00%	90,71%	92,06%	
97,21%	99,93%	100,00%	100,00%	70,00%	70,00%	63,64%	92,36%	92,36%	
95,74%	99,96%	100,00%	100,00%	60,00%	60,00%	50,00%	89,49%	89,49%	
97,15%	99,97%	100,00%	100,00%	80,00%	80,00%	66,67%	93,66%	93,66%	
96,06%	99,94%	100,00%	100,00%	66,67%	66,67%	50,00%	90,27%	90,27%	
99,42%	96,95%	83,33%	90,91%	63,64%	70,00%	63,64%	89,79%	91,18%	
98,75%	99,17%	50,00%	80,00%	14,29%	25,00%	71,43%	82,84%	86,92%	
99,09%	98,71%	77,78%	100,00%	44,44%	57,14%	44,44%	85,90%	89,39%	
98,10%	99,07%	100,00%	100,00%	100,00%	100,00%	66,67%	95,68%	95,68%	
98,63%	97,61%	87,50%	100,00%	50,00%	57,14%	50,00%	87,43%	89,40%	
99,26%	95,58%	85,71%	100,00%	71,43%	83,33%	62,50%	90,16%	92,78%	
99,67%	85,49%	100,00%	100,00%	62,50%	62,50%	55,56%	86,61%	86,61%	
98,64%	98,18%	91,67%	100,00%	75,00%	81,82%	75,00%	93,05%	94,57%	
97,64%	99,85%	90,91%	100,00%	63,64%	70,00%	69,23%	91,50%	93,04%	
99,04%	98,91%	88,89%	100,00%	66,67%	75,00%	66,67%	91,50%	93,45%	
97,85%	98,56%	78,69%	98,46%	51,04%	64,06%	61,79%	87,90%	91,18%	TOTAL

Tabela A.8 – Tabela de Resultados



Figura A.5 – Exemplo de Classificação

Vídeo: gsf-wm.2006.english.vs.paraguay.avi

Resolução: 576x320

Qualidade: BOM

CIM	CIA	RAIM	RAIA	TRS	RCC	TJ	JCA	JCC	JNS	
131465	132350	47372	50361	12	6	14	1	2	5	
128523	129849	51571	51825	16	8	18	2	8		
132558	132823	47928	49828	10	7	12	1	7	1	
131180	132028	49360	49871	20	9	19	1	9	1	
129709	130810	50866	51440	19	15	14	0	10		
172348	173512	3333	4744	11	5	9	1	5	1	
130142	130268	50173	52417	11	6	14	0	6	3	
126016	127180	52753	55584	9	5	20	1	5	11	
98780	99206	70636	83797	10	6	21	0	1	17	
117755	118966	60448	63626	13	11	22	1	2	18	
109885	110243	70508	73080	3	2	13	0	1	11	
128085	128452	52900	53583	15	10	18	0	10	4	
130659	131375	50573	51192	11	7	13	2	6	1	
132739	133849	45619	48361	13	9	10	0	7		
123183	124153	56554	57945	15	10	15	0	9	2	
105321	105773	69004	75652	18	13	19	3	9	4	
115603	116013	63447	65672	8	5	7	1	4		
99349	100183	79393	81958	14	13	16	4	11		
87328	87708	93305	95233	4	3	11	1	3	6	
114974	115084	64560	67042	14	8	15	0	8	1	
2445602	2459825	1130303	1183211	246	158	300	19	123	86	Total

Tabela A.9 – Tabela de dados

Vídeo: gsf-wm.2006.england.vs.paraguay.avi
 Resolução: 576x320
 Qualidade: BOM

CMPO	ADJ	JSC1	JSC2	JCC1	JCC2	RCCP	IQ1	IQ2	
99,33%	94,06%	57,14%	61,54%	22,22%	25,00%	50,00%	80,63%	81,34%	
98,98%	99,51%	88,89%	100,00%	44,44%	50,00%	50,00%	87,80%	89,47%	
99,80%	96,19%	83,33%	90,91%	63,64%	70,00%	70,00%	90,29%	91,69%	
99,36%	98,98%	89,47%	94,44%	50,00%	52,94%	45,00%	87,86%	88,66%	
99,16%	98,88%	100,00%	100,00%	71,43%	71,43%	78,95%	94,35%	94,35%	
99,33%	70,26%	77,78%	87,50%	62,50%	71,43%	45,45%	77,93%	79,79%	
99,90%	95,72%	78,57%	78,57%	54,55%	54,55%	54,55%	87,23%	87,23%	
99,08%	94,91%	40,00%	42,11%	55,56%	62,50%	55,56%	83,01%	83,91%	
99,57%	84,29%	19,05%	19,05%	25,00%	25,00%	60,00%	74,76%	74,76%	
98,98%	95,01%	13,64%	14,29%	50,00%	66,67%	84,62%	82,72%	84,45%	
99,68%	96,48%	15,38%	15,38%	50,00%	50,00%	66,67%	81,86%	81,86%	
99,71%	98,73%	77,78%	77,78%	71,43%	71,43%	66,67%	91,04%	91,04%	
99,45%	98,79%	76,92%	90,91%	50,00%	60,00%	63,64%	88,44%	90,84%	
99,17%	94,33%	100,00%	100,00%	70,00%	70,00%	69,23%	91,65%	91,65%	
99,22%	97,60%	86,67%	86,67%	69,23%	69,23%	66,67%	91,14%	91,14%	
99,57%	91,21%	63,16%	75,00%	60,00%	75,00%	72,22%	86,31%	89,00%	
99,65%	96,61%	85,71%	100,00%	57,14%	66,67%	62,50%	89,23%	91,61%	
99,17%	96,87%	75,00%	100,00%	68,75%	91,67%	92,86%	92,27%	97,07%	
99,57%	97,98%	36,36%	40,00%	60,00%	75,00%	75,00%	86,28%	88,14%	
99,90%	96,30%	93,33%	93,33%	57,14%	57,14%	57,14%	89,43%	89,43%	
99,42%	95,53%	65,00%	69,40%	57,48%	63,08%	64,23%	86,90%	87,90%	TOTAL

Tabela A.10 – Tabela de Resultados



Figura A.6 – Exemplo de Classificação

Vídeo: gsf-wm.2006-england.vs.ecuador.avi

Resolução: 512x384

Qualidade: BOM

CIM	CIA	RAIM	RAIA	TRS	RCC	TJ	JCA	JCC	JNS	
92599	94761	94667	100138	15	11	7	1	2	0	
137516	138818	52604	56091	15	7	9	0	2	0	
144405	146740	43899	47876	13	7	8	1	4	0	
143536	144987	47590	50131	8	7	9	1	7	0	
162943	164677	26452	29255	16	9	13	2	9	0	
130799	133505	59207	60841	10	6	7	0	2	0	
138855	140089	46862	54632	16	9	9	1	1	0	
107013	107754	85964	88072	6	5	7	1	5	0	
145957	147949	43875	46067	17	12	14	0	11	0	
155973	157357	34244	37129	13	9	12	1	9	0	
119415	121002	70382	73455	11	8	9	2	7	0	
132100	133803	51122	60878	16	11	15	0	10	0	
132969	134219	53517	60914	10	8	10	0	8	0	
157255	158889	29611	35008	10	7	10	0	7	0	
160075	162150	29901	31759	16	11	14	1	9	0	
172100	175832	9110	14263	12	7	11	1	7	0	
173299	176192	15165	18225	10	7	7	1	3	0	
138783	141229	50747	52876	10	7	6	2	2	0	
165071	167426	23469	27006	12	8	8	0	4	0	
157225	159024	34520	35735	9	6	10	1	6	0	
2867888	2906403	902908	980351	245	162	195	16	115	0	Total

Tabela A.11 – Tabela de dados

Vídeo: gsf-wm.2006-england.vs.ecuador.avi
 Resolução: 512x384
 Qualidade: BOM

CMPO	ADJ	JSC1	JSC2	JCC1	JCC2	RCCP	IQ1	IQ2	
97,72%	94,54%	85,71%	100,00%	28,57%	33,33%	73,33%	86,05%	87,96%	
99,06%	93,78%	100,00%	100,00%	22,22%	22,22%	46,67%	84,38%	84,38%	
98,41%	91,69%	87,50%	100,00%	50,00%	57,14%	53,85%	85,67%	87,63%	
99,00%	94,93%	88,89%	100,00%	77,78%	87,50%	87,50%	93,29%	95,38%	
98,95%	90,42%	84,62%	100,00%	69,23%	81,82%	56,25%	87,29%	90,08%	
97,97%	97,31%	100,00%	100,00%	28,57%	28,57%	60,00%	87,21%	87,21%	
99,12%	85,78%	88,89%	100,00%	11,11%	12,50%	56,25%	80,34%	81,59%	
99,31%	97,61%	85,71%	100,00%	71,43%	83,33%	83,33%	92,97%	95,59%	
98,65%	95,24%	100,00%	100,00%	78,57%	78,57%	70,59%	92,78%	92,78%	
99,12%	92,23%	91,67%	100,00%	75,00%	81,82%	69,23%	90,56%	92,08%	
98,69%	95,82%	77,78%	100,00%	77,78%	100,00%	72,73%	90,91%	95,35%	
98,73%	83,97%	100,00%	100,00%	66,67%	66,67%	68,75%	87,49%	87,49%	
99,07%	87,86%	100,00%	100,00%	80,00%	80,00%	80,00%	91,42%	91,42%	
98,97%	84,58%	100,00%	100,00%	70,00%	70,00%	70,00%	88,24%	88,24%	
98,72%	94,15%	92,86%	100,00%	64,29%	69,23%	68,75%	90,09%	91,30%	
97,88%	63,87%	90,91%	100,00%	63,64%	70,00%	58,33%	77,90%	79,45%	
98,36%	83,21%	85,71%	100,00%	42,86%	50,00%	70,00%	83,41%	85,55%	
98,27%	95,97%	66,67%	100,00%	33,33%	50,00%	70,00%	84,98%	89,98%	
98,59%	86,90%	100,00%	100,00%	50,00%	50,00%	66,67%	86,59%	86,59%	
98,87%	96,60%	90,00%	100,00%	60,00%	66,67%	66,67%	90,08%	91,75%	
98,67%	92,10%	91,79%	100,00%	58,97%	64,25%	66,12%	88,46%	89,81%	TOTAL

Tabela A.12 – Tabela de Resultados



Figura A.7 – Exemplo de Classificação

Vídeo: UEFA Cup SF 2nd L - Boro v Bucharest - 2nd Half.avi

Resolução: 640x368

Qualidade: BOM

CIM	CIA	RAIM	RAIA	TRS	RCC	TJ	JCA	JCC	JNS	
221015	223614	4771	5244	19	14	15	1	12	0	
227525	229549	0	0	11	9	14	4	9	0	
208956	210939	18449	20331	20	14	16	2	11	0	
179867	181186	51679	52709	9	7	7	1	6	0	
190678	192749	38689	40658	9	7	9	1	6	0	
206601	210707	18921	20072	8	4	5	0	4	0	
185291	186849	43899	45526	13	9	14	1	9	1	
171122	173256	57110	59377	10	5	6	0	5	0	
171011	172022	59821	61880	8	6	6	1	5	0	
192894	195069	36130	37221	17	12	12	0	6	1	
213550	217977	8362	9080	28	18	18	0	7	1	
225005	227997	0	0	13	10	12	1	10	0	
225320	228618	0	0	13	11	14	5	6	1	
211899	213105	17596	20193	7	7	6	0	6	0	
188489	190553	38706	41380	13	9	10	1	5	0	
220423	226022	3281	3490	17	11	9	0	3	0	
194840	198703	32005	33164	22	13	16	1	13	1	
210025	212821	17963	19533	16	14	16	1	13	0	
165172	166334	62443	65412	8	7	13	4	7	1	
188187	189922	42291	44282	6	5	5	2	1	0	
3997870	4047992	552116	579552	267	192	223	26	144	6	Total

Tabela A.13 – Tabela de dados

Vídeo: UEFA Cup SF 2nd L - Boro v Bucharest - 2nd Half.avi
 Resolução: 640x368
 Qualidade: BOM

CMPO	ADJ	JSC1	JSC2	JCC1	JCC2	RCCP	IQ1	IQ2	
98,84%	90,98%	93,33%	100,00%	80,00%	85,71%	73,68%	91,14%	92,38%	
99,12%	100,00%	71,43%	100,00%	64,29%	90,00%	81,82%	91,44%	96,87%	
99,06%	90,74%	87,50%	100,00%	68,75%	78,57%	70,00%	89,06%	91,29%	
99,27%	98,05%	85,71%	100,00%	85,71%	100,00%	77,78%	93,98%	96,84%	
98,93%	95,16%	88,89%	100,00%	66,67%	75,00%	77,78%	91,26%	93,21%	
98,05%	94,27%	100,00%	100,00%	80,00%	80,00%	50,00%	90,31%	90,31%	
99,17%	96,43%	85,71%	92,31%	69,23%	75,00%	69,23%	90,87%	92,11%	
98,77%	96,18%	100,00%	100,00%	83,33%	83,33%	50,00%	91,57%	91,57%	
99,41%	96,67%	83,33%	100,00%	83,33%	100,00%	75,00%	92,80%	96,13%	
98,89%	97,07%	91,67%	91,67%	54,55%	54,55%	70,59%	90,26%	90,26%	
97,97%	92,09%	94,44%	94,44%	41,18%	41,18%	64,29%	86,51%	86,51%	
98,69%	100,00%	91,67%	100,00%	83,33%	90,91%	76,92%	94,73%	96,32%	
98,56%	100,00%	57,14%	88,89%	46,15%	75,00%	84,62%	88,29%	94,35%	
99,43%	87,14%	100,00%	100,00%	100,00%	100,00%	100,00%	95,30%	95,30%	
98,92%	93,54%	90,00%	100,00%	50,00%	55,56%	69,23%	88,28%	89,84%	
97,52%	94,01%	100,00%	100,00%	33,33%	33,33%	64,71%	86,84%	86,84%	
98,06%	96,51%	87,50%	93,33%	86,67%	92,86%	59,09%	91,42%	92,62%	
98,69%	91,96%	93,75%	100,00%	81,25%	86,67%	87,50%	92,98%	94,14%	
99,30%	95,46%	61,54%	88,89%	58,33%	87,50%	87,50%	88,90%	94,56%	
99,09%	95,50%	60,00%	100,00%	20,00%	33,33%	83,33%	84,44%	89,77%	
98,76%	95,27%	85,65%	96,95%	66,36%	75,39%	71,91%	90,30%	92,34%	TOTAL

Tabela A.14 – Tabela de Resultados



Figura A.8 – Exemplo de Classificação

Vídeo: WC98 Brasil-France.avi

Resolução: 640x480

Qualidade: RUIM

CIM	CIA	RAIM	RAIA	TRS	RCC	TJ	JCA	JCC	JNS	
168941	172088	0	0	9	8	7	0	0	6	
173764	178320	112565	124822	10	8	9	3	0	3	
183007	186635	103801	117509	7	6	7	1	2	4	
240279	243638	55431	59327	12	6	12	1	6	4	
245946	252592	46610	48014	7	4	9	0	4	3	
249628	255967	38119	39667	12	8	13	0	8	4	
191938	194945	98731	107056	18	12	15	2	5	6	
183702	186870	91708	116992	14	14	4	1	0	3	
145015	148447	123271	156088	18	14	6	2	0	0	
237101	243163	51543	59206	6	4	8	2	1	5	
230887	236585	56991	62400	14	8	12	0	3	2	
241627	248086	43725	49435	11	8	11	1	4	4	
184602	192602	102082	109142	10	6	10	4	1	1	
129782	134413	146036	168449	18	14	8	3	1	1	
234566	237909	46496	60077	18	11	10	2	0	1	
185582	193472	106845	112618	5	5	2	0	0	2	
178540	182975	101125	121783	7	5	3	0	2	1	
222077	225901	67362	78886	8	7	6	0	2	3	
156746	161114	127389	140811	5	5	14	1	2	11	
183583	188394	96654	112676	14	11	14	0	11	0	
3967313	4064116	1616484	1844958	223	164	180	23	52	64	Total

Tabela A.15 – Tabela de dados

Vídeo: WC98 Brasil-France.avi
 Resolução: 640x480
 Qualidade: RUIM

CMPO	ADJ	JSC1	JSC2	JCC1	JCC2	RCCP	IQ1	IQ2	
98,17%	100,00%	14,29%	14,29%	0,00%	0,00%	88,89%	79,68%	79,68%	
97,45%	90,18%	33,33%	50,00%	0,00%	0,00%	80,00%	77,00%	78,67%	
98,06%	88,33%	28,57%	33,33%	66,67%	100,00%	85,71%	83,33%	87,14%	
98,62%	93,43%	58,33%	63,64%	75,00%	85,71%	50,00%	85,55%	87,15%	
97,37%	97,08%	66,67%	66,67%	66,67%	66,67%	57,14%	87,10%	87,10%	
97,52%	96,10%	69,23%	69,23%	88,89%	88,89%	66,67%	90,25%	90,25%	
98,46%	92,22%	46,67%	53,85%	55,56%	71,43%	66,67%	83,63%	85,93%	
98,30%	78,39%	0,00%	0,00%	0,00%	#DIV/0!	100,00%	71,84%	#DIV/0!	
97,69%	78,98%	66,67%	100,00%	0,00%	0,00%	77,78%	76,28%	79,61%	
97,51%	87,06%	12,50%	16,67%	33,33%	100,00%	66,67%	75,85%	82,93%	
97,59%	91,33%	83,33%	83,33%	30,00%	30,00%	57,14%	83,17%	83,17%	
97,40%	88,45%	54,55%	60,00%	57,14%	66,67%	72,73%	83,49%	84,99%	
95,85%	93,53%	50,00%	83,33%	11,11%	20,00%	60,00%	78,39%	82,62%	
96,55%	86,69%	50,00%	80,00%	14,29%	25,00%	77,78%	78,34%	82,41%	
98,59%	77,39%	70,00%	87,50%	0,00%	0,00%	61,11%	74,71%	76,46%	
95,92%	94,87%	0,00%	0,00%	#DIV/0!	#DIV/0!	100,00%	#DIV/0!	#DIV/0!	
97,58%	83,04%	66,67%	66,67%	100,00%	100,00%	71,43%	87,02%	87,02%	
98,31%	85,39%	50,00%	50,00%	66,67%	66,67%	87,50%	84,71%	84,71%	
97,29%	90,47%	14,29%	15,38%	66,67%	100,00%	100,00%	83,81%	87,25%	
97,45%	85,78%	100,00%	100,00%	78,57%	78,57%	78,57%	89,84%	89,84%	
97,62%	87,62%	51,67%	59,24%	44,83%	55,91%	73,54%	81,84%	83,70%	TOTAL

Tabela A.16 – Tabela de Resultados



Figura A.9 – Exemplo de Classificação

Vídeo: World.Cup.2006.Group.A.Poland.vs.Ecuador.HR.HDTV.CD2.avi

Resolução: 960x544

Qualidade: ÓTIMO

CIM	CIA	RAIM	RAIA	TRS	RCC	TJ	JCA	JCC	JNS	
347048	352401	127048	162379	12	9	11	1	5	0	
370441	375314	102206	138765	10	6	11	1	0	0	
351875	354409	120550	164337	30	24	9	1	3	0	
391044	394409	93792	121687	18	13	9	1	3	0	
387148	390616	83981	124800	19	8	14	1	7	0	
382980	392697	89112	122774	17	11	11	4	1	0	
389747	392602	83292	122582	21	13	14	1	7	0	
371078	373805	115249	144163	18	9	10	0	7	0	
322134	324703	156471	193742	5	4	9	1	0	7	
286988	289279	190615	228441	14	10	10	1	7	0	
262610	265629	212859	250362	15	6	16	7	5	0	
257649	261197	218709	255645	13	10	13	5	5	0	
389889	393578	82413	121416	20	12	15	4	6	0	
390686	394895	94234	120693	31	23	12	1	3	0	
280271	285094	190518	230182	16	6	16	0	5	5	
356981	362507	111885	152255	10	5	15	0	5	8	
292151	300858	169714	213519	31	27	6	1	1	0	
369703	372594	107774	144158	36	29	8	0	2	0	
380056	384244	82575	129714	14	10	15	2	3	7	
390767	393094	85401	123491	13	12	11	2	2	6	
6971246	7053925	2518398	3265105	363	247	235	34	77	33	Total

Tabela A.17 – Tabela de dados

Vídeo: World.Cup.2006.Group.A.Poland.vs.Ecuador.HR.HDTV.CD2.avi
 Resolução: 960x544
 Qualidade: ÓTIMO

CMPO	ADJ	JSC1	JSC2	JCC1	JCC2	RCCP	IQ1	IQ2	
98,48%	78,24%	90,91%	100,00%	45,45%	50,00%	75,00%	82,99%	84,35%	
98,70%	73,65%	90,91%	100,00%	0,00%	0,00%	60,00%	75,42%	76,32%	
99,29%	73,36%	88,89%	100,00%	33,33%	37,50%	80,00%	80,65%	82,17%	
99,15%	77,08%	88,89%	100,00%	33,33%	37,50%	72,22%	81,12%	82,65%	
99,11%	67,29%	92,86%	100,00%	50,00%	53,85%	42,11%	76,74%	77,84%	
97,53%	72,58%	63,64%	100,00%	9,09%	14,29%	64,71%	73,28%	77,44%	
99,27%	67,95%	92,86%	100,00%	50,00%	53,85%	61,90%	79,00%	80,10%	
99,27%	79,94%	100,00%	100,00%	70,00%	70,00%	50,00%	84,72%	84,72%	
99,21%	80,76%	11,11%	12,50%	0,00%	0,00%	80,00%	72,10%	72,24%	
99,21%	83,44%	90,00%	100,00%	70,00%	77,78%	71,43%	87,07%	88,85%	
98,86%	85,02%	56,25%	100,00%	31,25%	55,56%	40,00%	77,11%	83,91%	
98,64%	85,55%	61,54%	100,00%	38,46%	62,50%	76,92%	82,16%	88,41%	
99,06%	67,88%	73,33%	100,00%	40,00%	54,55%	60,00%	75,76%	79,88%	
98,93%	78,08%	91,67%	100,00%	25,00%	27,27%	74,19%	81,04%	82,10%	
98,31%	82,77%	68,75%	68,75%	45,45%	45,45%	37,50%	78,55%	78,55%	
98,48%	73,49%	46,67%	46,67%	71,43%	71,43%	50,00%	77,00%	77,00%	
97,11%	79,48%	83,33%	100,00%	16,67%	20,00%	87,10%	80,52%	82,52%	
99,22%	74,76%	100,00%	100,00%	25,00%	25,00%	80,56%	81,45%	81,45%	
98,91%	63,66%	40,00%	46,15%	37,50%	50,00%	71,43%	71,79%	73,66%	
99,41%	69,16%	27,27%	33,33%	40,00%	66,67%	92,31%	74,96%	78,23%	
98,83%	77,13%	71,49%	83,58%	38,12%	45,83%	68,04%	79,35%	81,33%	TOTAL

Tabela A.18 – Tabela de Resultados



Figura A.10 – Exemplo de Classificação

Vídeo: World.Cup.2006.Group.B.Trinidad.and.Tobago.vs.Sweden.HR.HDTV.avi

Resolução: 960x544

Qualidade: ÓTIMO

CIM	CIA	RAIM	RAIA	TRS	RCC	TJ	JCA	JCC	JNS	
312282	333513	178237	182260	24	22	7	3	3	1	
366718	384106	121649	133950	32	30	9	2	5	1	
429132	438604	79509	81859	22	21	5	0	4	0	
429680	438154	74040	77905	24	23	12	0	11	0	
388600	404117	105428	113505	30	27	12	1	8	0	
364102	421889	86778	90187	31	25	12	1	5	0	
437092	446890	64413	67270	28	24	15	1	10	0	
438236	443048	72209	76037	35	33	7	2	3	0	
412323	431540	79841	82232	29	27	13	2	11	0	
434690	445903	63835	68334	29	13	14	1	7	1	
435116	450941	61039	63382	29	21	16	1	9	1	
440809	447818	65092	69739	27	19	12	0	5	1	
442945	448490	63671	68180	26	21	14	2	8	0	
438728	447324	64218	68878	30	26	14	1	10	0	
436678	444722	65034	69754	35	31	15	1	11	0	
417770	434305	78385	81504	41	34	18	0	11	1	
333180	354582	154697	164725	28	22	11	1	4	1	
304684	368955	144935	146745	34	29	9	1	5	3	
319079	335366	169220	181537	25	19	12	4	6	0	
344922	377466	128692	139719	38	33	12	1	8	0	
7926766	8297733	1920922	2027702	597	500	239	25	144	10	Total

Tabela A.19 – Tabela de dados

Vídeo: World.Cup.2006.Group.B.Trinidad.and.Tobago.vs.Sweden.HR.HDTV.avi
 Resolução: 960x544
 Qualidade: ÓTIMO

CMPO	ADJ	JSC1	JSC2	JCC1	JCC2	RCCP	IQ1	IQ2	
93,63%	97,79%	42,86%	75,00%	50,00%	100,00%	91,67%	85,45%	93,67%	
95,47%	90,82%	66,67%	85,71%	62,50%	83,33%	93,75%	87,49%	91,48%	
97,84%	97,13%	100,00%	100,00%	80,00%	80,00%	95,45%	95,78%	95,78%	
98,07%	95,04%	100,00%	100,00%	91,67%	91,67%	95,83%	96,34%	96,34%	
96,16%	92,88%	91,67%	100,00%	66,67%	72,73%	90,00%	91,00%	92,44%	
86,30%	96,22%	91,67%	100,00%	41,67%	45,45%	80,65%	85,28%	86,49%	
97,81%	95,75%	93,33%	100,00%	66,67%	71,43%	85,71%	92,32%	93,46%	
98,91%	94,97%	71,43%	100,00%	42,86%	60,00%	94,29%	88,71%	93,29%	
95,55%	97,09%	84,62%	100,00%	84,62%	100,00%	93,10%	93,66%	96,73%	
97,49%	93,42%	85,71%	92,31%	53,85%	58,33%	44,83%	85,25%	86,36%	
96,49%	96,30%	87,50%	93,33%	60,00%	64,29%	72,41%	89,47%	90,48%	
98,43%	93,34%	91,67%	91,67%	45,45%	45,45%	70,37%	87,87%	87,87%	
98,76%	93,39%	85,71%	100,00%	57,14%	66,67%	80,77%	89,62%	92,00%	
98,08%	93,23%	92,86%	100,00%	71,43%	76,92%	86,67%	92,05%	93,32%	
98,19%	93,23%	93,33%	100,00%	73,33%	78,57%	88,57%	92,52%	93,71%	
96,19%	96,17%	94,44%	94,44%	64,71%	64,71%	82,93%	91,54%	91,54%	
93,96%	93,91%	81,82%	90,00%	40,00%	44,44%	78,57%	85,80%	87,06%	
82,58%	98,77%	55,56%	62,50%	83,33%	100,00%	85,29%	85,89%	88,25%	
95,14%	93,22%	66,67%	100,00%	50,00%	75,00%	76,00%	85,19%	91,03%	
91,38%	92,11%	91,67%	100,00%	66,67%	72,73%	86,84%	88,74%	90,18%	
95,53%	94,73%	85,36%	95,33%	62,88%	70,59%	83,75%	89,79%	91,56%	TOTAL

Tabela A.20 – Tabela de Resultados



Figura A.11 – Exemplo de Classificação

Vídeo: Copa2006.BRASILxAUSTRALIA.1oTempo.Portugues.xara.avi

Resolução: 672x368

Qualidade: ÓTIMO

CIM	CIA	RAIM	RAIA	TRS	RCC	TJ	JCA	JCC	JNS	
243363	243660	0	0	15	10	21	1	10	5	
220053	220824	18639	24161	10	6	13	2	6	3	
191941	192729	31763	52112	16	12	12	1	2	4	
833	148568	31065	97440	0	0	7	0	0	0	
212730	214338	26616	30256	14	6	12	0	6	0	
186989	187781	54051	57560	7	7	9	1	7	1	
224811	226069	10222	17941	12	5	11	0	4	1	
164540	168075	69890	74435	7	5	8	2	4	0	
214308	214953	26268	30869	8	7	8	0	4	1	
207731	208440	33168	37784	4	3	6	0	2	2	
212911	213425	23051	31844	9	6	10	0	6	2	
237213	237503	0	0	16	9	8	0	3	1	
225793	225939	13186	19835	8	5	10	0	4	4	
224593	224942	13005	19388	14	7	10	0	7	1	
201849	202542	34933	41270	12	9	12	1	9	0	
195819	197579	44683	47507	11	8	10	0	7	0	
191094	192422	50203	52639	10	7	10	0	7	0	
137763	140279	95837	104358	14	8	11	0	8	1	
211912	212392	28952	32722	12	10	15	1	10	3	
227792	228369	7003	8338	15	11	15	1	9	3	
3934038	4100829	612535	780459	214	141	218	10	115	32	Total

Tabela A.21 – Tabela de dados

Vídeo: Copa2006.BRASILxAUSTRALIA.1oTempo.Portugues.xara.avi
 Resolução: 672x368
 Qualidade: ÓTIMO

CMPO	ADJ	JSC1	JSC2	JCC1	JCC2	RCCP	IQ1	IQ2	
99,88%	100,00%	71,43%	75,00%	62,50%	66,67%	66,67%	90,02%	90,79%	
99,65%	77,14%	61,54%	72,73%	60,00%	75,00%	60,00%	80,03%	82,65%	
99,59%	60,95%	58,33%	63,64%	25,00%	28,57%	75,00%	72,02%	72,91%	
0,56%	31,88%	100,00%	100,00%	0,00%	0,00%	#DIV/0!	#DIV/0!	#DIV/0!	
99,25%	87,97%	100,00%	100,00%	50,00%	50,00%	42,86%	84,81%	84,81%	
99,58%	93,90%	77,78%	87,50%	87,50%	100,00%	100,00%	94,25%	96,47%	
99,44%	56,98%	90,91%	90,91%	40,00%	40,00%	41,67%	72,00%	72,00%	
97,90%	93,89%	75,00%	100,00%	50,00%	66,67%	71,43%	86,77%	90,94%	
99,70%	85,10%	87,50%	87,50%	57,14%	57,14%	87,50%	87,89%	87,89%	
99,66%	87,78%	66,67%	66,67%	50,00%	50,00%	75,00%	84,77%	84,77%	
99,76%	72,39%	80,00%	80,00%	75,00%	75,00%	66,67%	82,42%	82,42%	
99,88%	100,00%	87,50%	87,50%	42,86%	42,86%	56,25%	88,62%	88,62%	
99,94%	66,48%	60,00%	60,00%	66,67%	66,67%	62,50%	77,16%	77,16%	
99,84%	67,08%	90,00%	90,00%	77,78%	77,78%	50,00%	80,20%	80,20%	
99,66%	84,65%	91,67%	100,00%	75,00%	81,82%	75,00%	88,67%	90,19%	
99,11%	94,06%	100,00%	100,00%	70,00%	70,00%	72,73%	91,88%	91,88%	
99,31%	95,37%	100,00%	100,00%	70,00%	70,00%	70,00%	92,14%	92,14%	
98,21%	91,83%	90,91%	90,91%	80,00%	80,00%	57,14%	89,32%	89,32%	
99,77%	88,48%	73,33%	78,57%	83,33%	90,91%	83,33%	89,89%	91,17%	
99,75%	83,99%	73,33%	78,57%	75,00%	81,82%	73,33%	86,47%	87,68%	
95,93%	78,48%	80,73%	84,62%	61,83%	65,34%	65,89%	81,89%	82,63%	TOTAL

Tabela A.22 – Tabela de Resultados



Figura A.12 – Exemplo de Classificação

Vídeo: World.Cup.2006.Group.A.Germany.vs.Costa.Rica.HR.HDTV.-CD1.avi

Resolução: 960x544

Qualidade: ÓTIMO

CIM	CIA	RAIM	RAIA	TRS	RCC	TJ	JCA	JCC	JNS	
328547	330088	179085	188558	14	12	20	2	12	4	
405622	407581	103918	107769	19	13	19	0	13	0	
422816	424291	89887	93031	15	11	16	0	11	1	
312711	317365	187696	196682	13	9	20	0	9	7	
245953	249513	247240	268362	11	9	21	0	9	10	
297875	300461	185563	216748	6	6	7	1	6	0	
375509	376690	135304	141845	11	7	11	1	7	1	
431575	432845	69616	76727	20	13	19	0	13	2	
443439	482298	733	733	17	13	19	1	13	1	
441760	443195	63031	64314	22	14	20	1	13	0	
358770	363080	139234	146375	16	15	15	2	12	0	
322711	327329	175643	182781	11	9	12	2	9	0	
355352	357082	147978	158629	10	7	13	3	7	1	
338330	344531	157824	162406	14	13	15	3	12	0	
364054	364808	145746	153446	15	8	17	1	7	2	
422318	423446	87453	93003	19	10	18	0	10	0	
412844	417997	83467	85412	16	15	17	1	15	0	
421295	422481	90051	95204	10	9	11	0	9	1	
278416	283868	199070	231281	4	2	3	0	2	0	
274489	279648	200067	239125	11	10	19	0	10	8	
7254386	7348597	2688606	2902431	274	205	312	18	199	38	Total

Tabela A.23 – Tabela de dados

Vídeo: World.Cup.2006.Group.A.Germany.vs.Costa.Rica.HR.HDTV.-CD1.avi
 Resolução: 960x544
 Qualidade: ÓTIMO

CMPO	ADJ	JSC1	JSC2	JCC1	JCC2	RCCP	IQ1	IQ2	
99,53%	94,98%	70,00%	77,78%	75,00%	85,71%	85,71%	91,15%	93,00%	
99,52%	96,43%	100,00%	100,00%	68,42%	68,42%	68,42%	92,27%	92,27%	
99,65%	96,62%	93,75%	93,75%	73,33%	73,33%	73,33%	92,74%	92,74%	
98,53%	95,43%	65,00%	65,00%	69,23%	69,23%	69,23%	88,23%	88,23%	
98,57%	92,13%	52,38%	52,38%	81,82%	81,82%	81,82%	88,35%	88,35%	
99,14%	85,61%	85,71%	100,00%	85,71%	100,00%	100,00%	91,81%	94,66%	
99,69%	95,39%	81,82%	90,00%	70,00%	77,78%	63,64%	89,82%	91,42%	
99,71%	90,73%	89,47%	89,47%	76,47%	76,47%	65,00%	89,75%	89,75%	
91,94%	100,00%	89,47%	94,44%	72,22%	76,47%	76,47%	91,00%	91,92%	
99,68%	98,01%	95,00%	100,00%	65,00%	68,42%	63,64%	91,55%	92,39%	
98,81%	95,12%	86,67%	100,00%	80,00%	92,31%	93,75%	93,92%	96,48%	
98,59%	96,09%	83,33%	100,00%	75,00%	90,00%	81,82%	92,15%	95,32%	
99,52%	93,29%	69,23%	90,00%	58,33%	77,78%	70,00%	87,24%	91,26%	
98,20%	97,18%	80,00%	100,00%	80,00%	100,00%	92,86%	93,67%	97,67%	
99,79%	94,98%	82,35%	87,50%	46,67%	50,00%	53,33%	86,41%	87,25%	
99,73%	94,03%	100,00%	100,00%	55,56%	55,56%	52,63%	88,64%	88,64%	
98,77%	97,72%	94,12%	100,00%	88,24%	93,75%	93,75%	96,38%	97,52%	
99,72%	94,59%	90,91%	90,91%	90,00%	90,00%	90,00%	95,10%	95,10%	
98,08%	86,07%	100,00%	100,00%	66,67%	66,67%	50,00%	86,12%	86,12%	
98,16%	83,67%	57,89%	57,89%	90,91%	90,91%	90,91%	87,61%	87,61%	
98,72%	92,63%	82,05%	87,07%	72,63%	77,73%	74,82%	89,92%	90,94%	TOTAL

Tabela A.24 – Tabela de Resultados



Figura A.13 – Exemplo de Classificação

Vídeo: germany vs brazil 2nd half.avi

Resolução: 512x304

Qualidade: RUIM

CIM	CIA	RAIM	RAIA	TRS	RCC	TJ	JCA	JCC	JNS	
97208	97791	48932	56811	4	3	5	0	1	1	
125196	128818	24012	24700	6	4	7	2	4		
66638	116439	32383	37784	3	2	6	2	2	1	
64176	115514	33789	38739	2	2	6	2	2	2	
50754	125519	24170	28050	4	3	12	3	3	5	
113371	114187	36405	40307	4	4	8	0	4	4	
67523	115030	35092	39106	7	6	11	0	4	6	
52732	104975	42831	47948	7	6	14	4	6	4	
46074	104997	45053	48640	5	3	15	4	2	5	
113319	116335	34909	38221	5	3	6	1	3		
125245	127891	24142	26915	3	3	4	0	3	1	
93223	117768	23180	34846	4	3	5	1	3	1	
105917	109079	35322	42730	10	6	11	2	6		
102078	110244	36050	42562	15	14	11	0	14	2	
96244	111638	37147	42356	10	10	10	2	6	2	
107788	108857	39476	45056	9	7	14	1	7	4	
72363	114603	34765	39549	6	4	10	1	4	2	
117600	118768	31551	35851	6	6	9	0	6	3	
118835	119745	30355	34039	7	6	11	1	6	3	
62931	108235	40501	45044	7	7	14	3	7	4	
1799215	2286433	690065	789254	124	102	189	29	93	50	Total

Tabela A.25 – Tabela de dados

Vídeo: germany vs brazil 2nd half.avi
 Resolução: 512x304
 Qualidade: RUIM

CMPO	ADJ	JSC1	JSC2	JCC1	JCC2	RCCP	IQ1	IQ2	
99,40%	86,13%	80,00%	80,00%	25,00%	25,00%	75,00%	82,94%	82,94%	
97,19%	97,21%	71,43%	100,00%	57,14%	80,00%	66,67%	87,56%	92,71%	
57,23%	85,71%	50,00%	75,00%	40,00%	66,67%	66,67%	65,69%	70,86%	
55,56%	87,22%	33,33%	50,00%	50,00%	100,00%	100,00%	68,31%	74,97%	
40,44%	86,17%	33,33%	44,44%	42,86%	75,00%	75,00%	59,43%	63,76%	
99,29%	90,32%	50,00%	50,00%	100,00%	100,00%	100,00%	91,36%	91,36%	
58,70%	89,74%	45,45%	45,45%	80,00%	80,00%	85,71%	73,07%	73,07%	
50,23%	89,33%	42,86%	60,00%	60,00%	100,00%	85,71%	67,70%	73,42%	
43,88%	92,63%	40,00%	54,55%	20,00%	33,33%	60,00%	59,78%	62,57%	
97,41%	91,33%	83,33%	100,00%	50,00%	60,00%	60,00%	85,39%	88,06%	
97,93%	89,70%	75,00%	75,00%	100,00%	100,00%	100,00%	93,17%	93,17%	
79,16%	66,52%	60,00%	75,00%	75,00%	100,00%	75,00%	71,99%	75,99%	
97,10%	82,66%	81,82%	100,00%	54,55%	66,67%	60,00%	82,55%	85,58%	
92,59%	84,70%	81,82%	81,82%	155,56%	155,56%	93,33%	95,12%	95,12%	
86,21%	87,70%	60,00%	75,00%	75,00%	100,00%	100,00%	84,37%	88,37%	
99,02%	87,62%	64,29%	69,23%	70,00%	77,78%	77,78%	86,53%	87,80%	
63,14%	87,90%	70,00%	77,78%	50,00%	57,14%	66,67%	71,53%	73,02%	
99,02%	88,01%	66,67%	66,67%	100,00%	100,00%	100,00%	92,12%	92,12%	
99,24%	89,18%	63,64%	70,00%	75,00%	85,71%	85,71%	88,38%	90,09%	
58,14%	89,91%	50,00%	63,64%	70,00%	100,00%	100,00%	73,82%	78,18%	
78,69%	87,43%	58,20%	68,75%	66,91%	84,55%	82,26%	78,88%	81,70%	TOTAL

Tabela A.26 – Tabela de Resultados



Figura A.14 – Exemplo de Classificação

Vídeo: world.cup.2006.group.d.iran.vs.mexico.hr.hdtv.ac3.5.1.xvid.avi

Resolução: 960x544

Qualidade: ÓTIMO

CIM	CIA	RAIM	RAIA	TRS	RCC	TJ	JCA	JCC	JNS	
190656	398192	101818	119040	8	5	15	0	5	7	
203281	386321	116804	130560	9	7	14	1	7	4	
260010	393345	105592	123840	9	8	11	1	8	1	
242592	399593	104061	118713	12	9	13	0	7	2	
257641	424037	83101	89857	12	10	16	2	10	2	
268048	408238	90332	108998	10	8	12	1	8	2	
286280	380000	117252	138231	9	8	12	1	8	2	
285181	359899	132138	154763	12	9	14	2	9	0	
264451	363404	139234	153980	13	9	15	2	8	1	
242869	419633	84393	98150	9	7	12	0	7	4	
212406	428066	78085	87617	11	8	15	0	8	4	
187256	380570	126094	138185	9	8	13	0	6	6	
221418	366534	133401	152205	8	6	12	2	6	2	
327322	344502	149890	170880	11	9	10	1	7	0	
219853	404212	102965	114260	10	6	14	2	5	4	
243913	412826	91638	105850	9	8	11	1	7	2	
234304	417450	83673	100137	8	8	10	1	7	2	
225853	430735	75565	79554	8	5	13	1	4	4	
273501	388843	107874	127902	28	19	17	1	7	0	
353087	369744	126862	145370	9	8	12	3	7	0	
4999922	7876144	2150772	2458092	214	165	261	22	141	49	Total

Tabela A.27 – Tabela de dados

Vídeo: world.cup.2006.group.d.iran.vs.mexico.hr.hdtv.ac3.5.1.xvid.avi
 Resolução: 960x544
 Qualidade: ÓTIMO

CMPO	ADJ	JSC1	JSC2	JCC1	JCC2	RCCP	IQ1	IQ2	
47,88%	85,53%	53,33%	53,33%	62,50%	62,50%	62,50%	64,53%	64,53%	
52,62%	89,46%	64,29%	69,23%	70,00%	77,78%	77,78%	70,94%	72,21%	
66,10%	85,26%	81,82%	90,00%	80,00%	88,89%	88,89%	78,05%	79,76%	
60,71%	87,66%	84,62%	84,62%	63,64%	63,64%	75,00%	74,25%	74,25%	
60,76%	92,48%	75,00%	85,71%	71,43%	83,33%	83,33%	76,61%	78,87%	
65,66%	82,87%	75,00%	81,82%	80,00%	88,89%	80,00%	75,49%	77,06%	
75,34%	84,82%	75,00%	81,82%	80,00%	88,89%	88,89%	80,44%	82,02%	
79,24%	85,38%	85,71%	100,00%	64,29%	75,00%	75,00%	80,12%	82,62%	
72,77%	90,42%	80,00%	92,31%	57,14%	66,67%	69,23%	77,76%	79,94%	
57,88%	85,98%	66,67%	66,67%	87,50%	87,50%	77,78%	73,55%	73,55%	
49,62%	89,12%	73,33%	73,33%	72,73%	72,73%	72,73%	70,44%	70,44%	
49,20%	91,25%	53,85%	53,85%	85,71%	85,71%	88,89%	72,00%	72,00%	
60,41%	87,65%	66,67%	80,00%	60,00%	75,00%	75,00%	71,99%	74,82%	
95,01%	87,72%	90,00%	100,00%	70,00%	77,78%	81,82%	88,14%	89,91%	
54,39%	90,11%	57,14%	66,67%	50,00%	62,50%	60,00%	67,29%	69,49%	
59,08%	86,57%	72,73%	80,00%	77,78%	87,50%	88,89%	74,92%	76,62%	
56,13%	83,56%	70,00%	77,78%	87,50%	100,00%	100,00%	74,64%	76,67%	
52,43%	94,99%	61,54%	66,67%	44,44%	50,00%	62,50%	68,45%	69,51%	
70,34%	84,34%	94,12%	100,00%	41,18%	43,75%	67,86%	74,45%	75,30%	
95,49%	87,27%	75,00%	100,00%	58,33%	77,78%	88,89%	86,19%	90,63%	
63,48%	87,50%	72,80%	79,50%	66,51%	74,21%	77,10%	74,48%	75,92%	TOTAL

Tabela A.28 – Tabela de Resultados

Apêndice B

Exemplos de imagens segmentadas e classificadas

As figuras a seguir são ilustrações do processo de segmentação automática e classificação em imagens capturadas em vários jogos, onde cada imagem da figura representa segundo a legenda:

(a)	(b)
(c)	(d)
(e)	

(a) – Imagem original

(b) – Marcação manual

(c) – Limiarização da imagem original, com os limiares obtidos na detecção do campo

(d) – Diferenças de marcação manual/automática, onde as tonalidades escuras de verde e vermelho representam marcações coincidentes, e as claras, de não coincidentes.

(e) – Imagem original com marcações de classificação

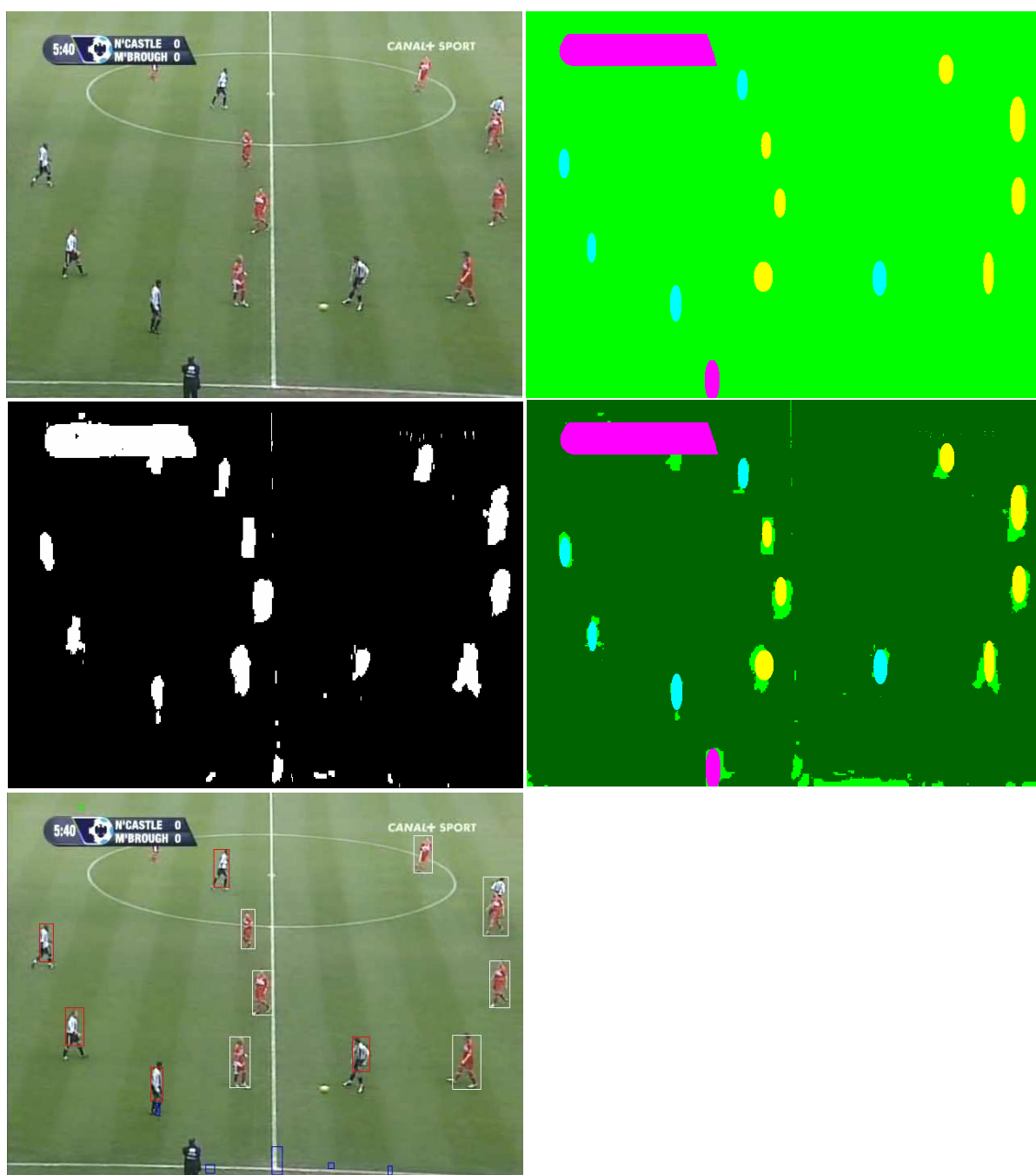


Figura B.1 – Exemplo

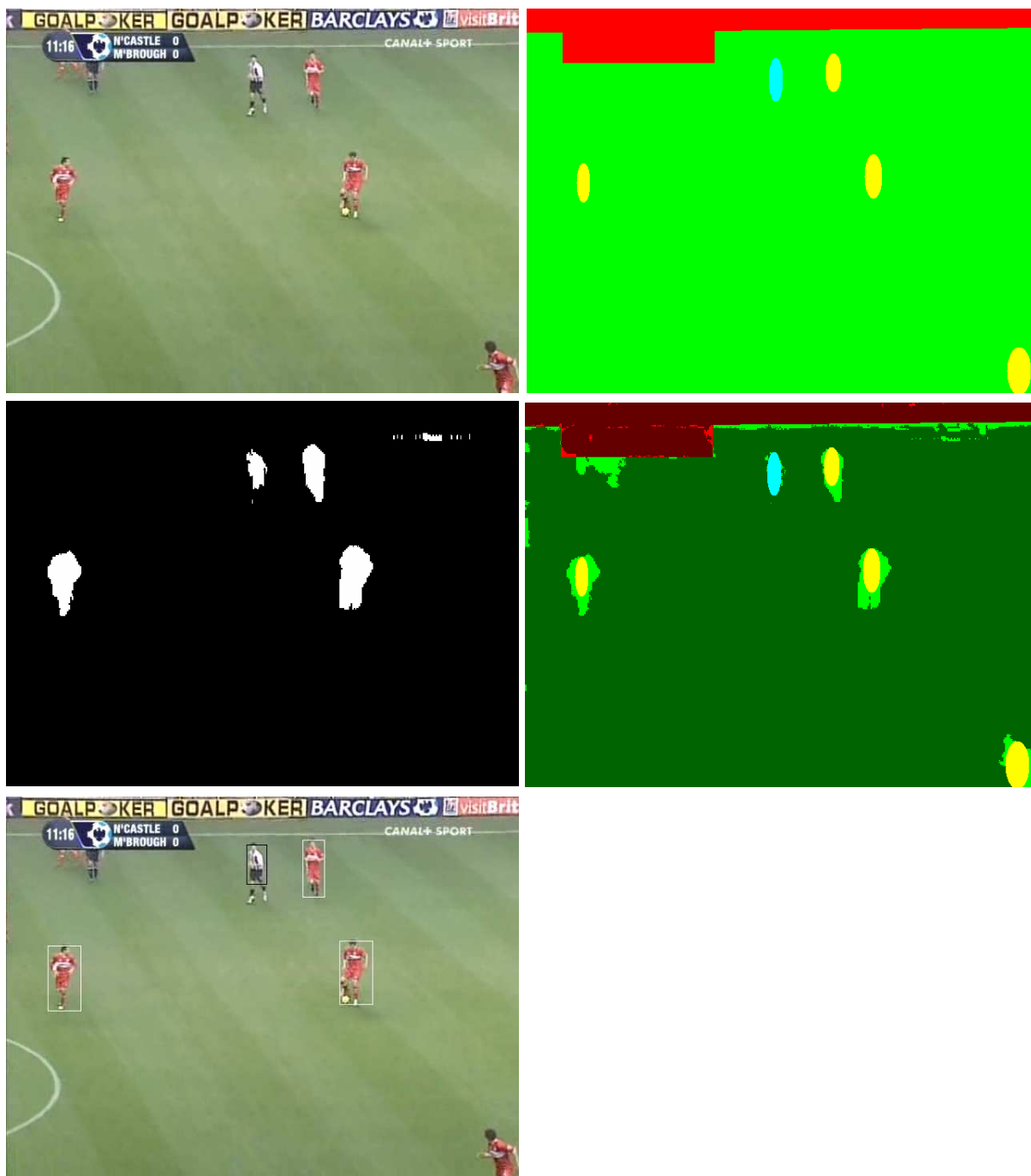


Figura B.2 – Exemplo

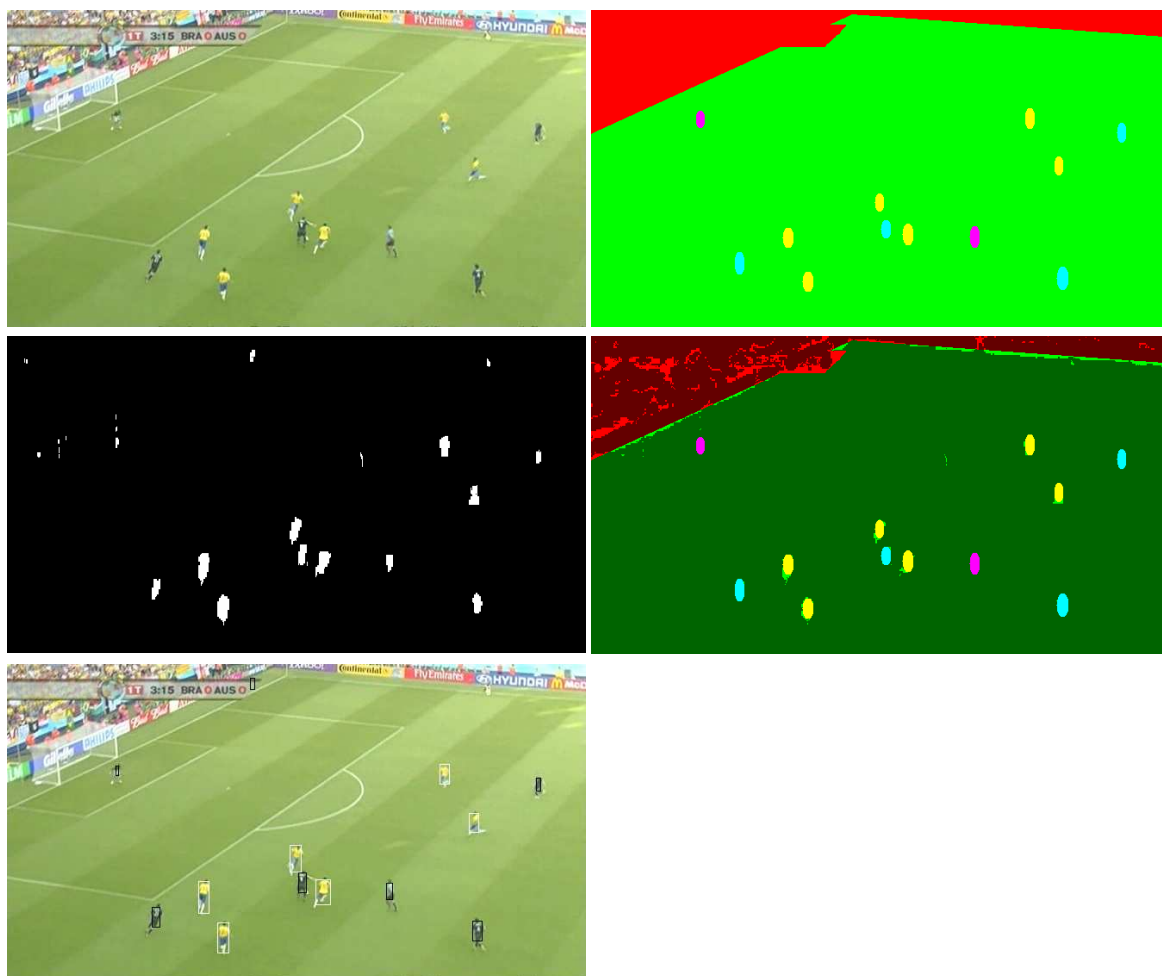


Figura B.3 – Exemplo

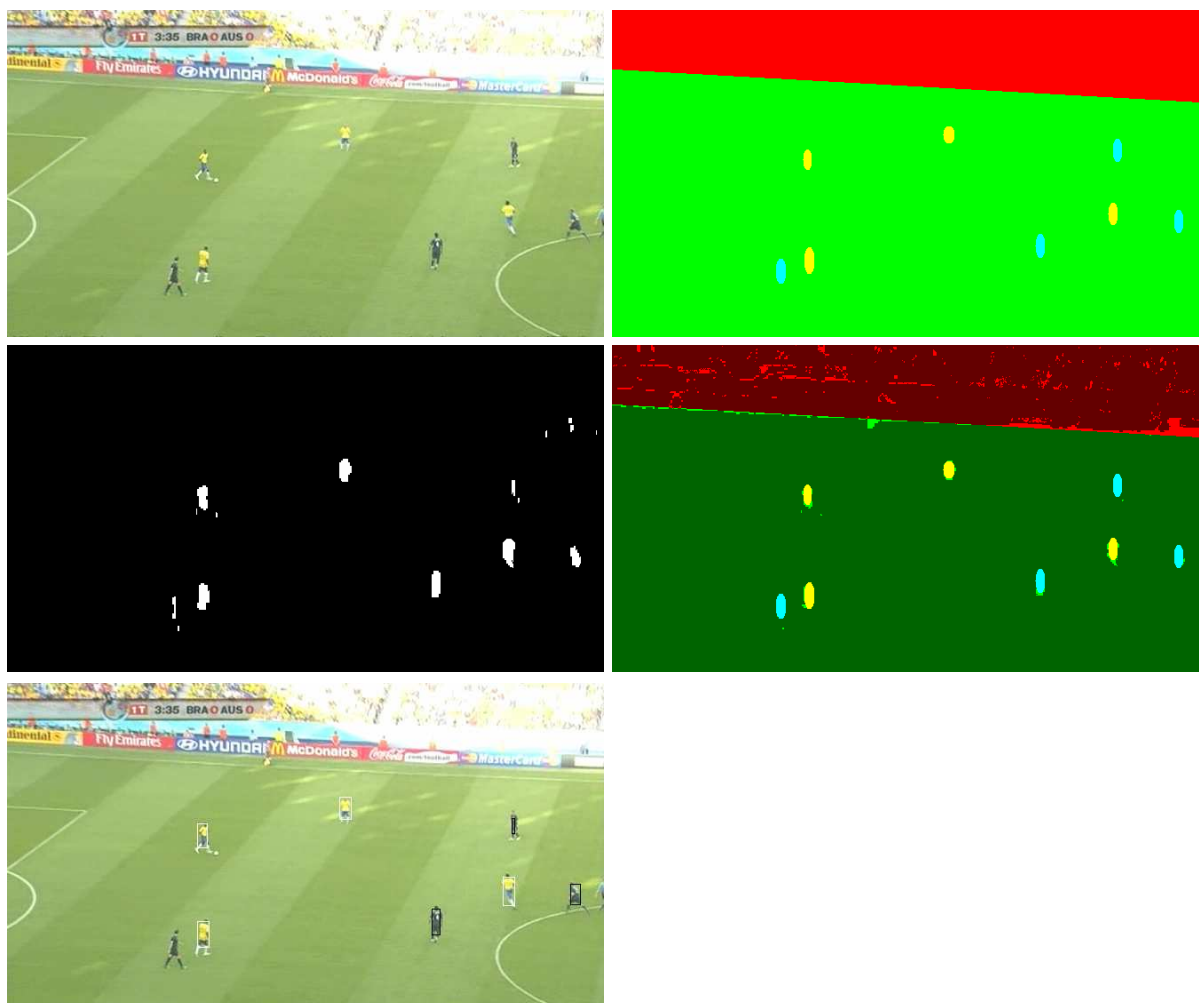


Figura B.4 – Exemplo

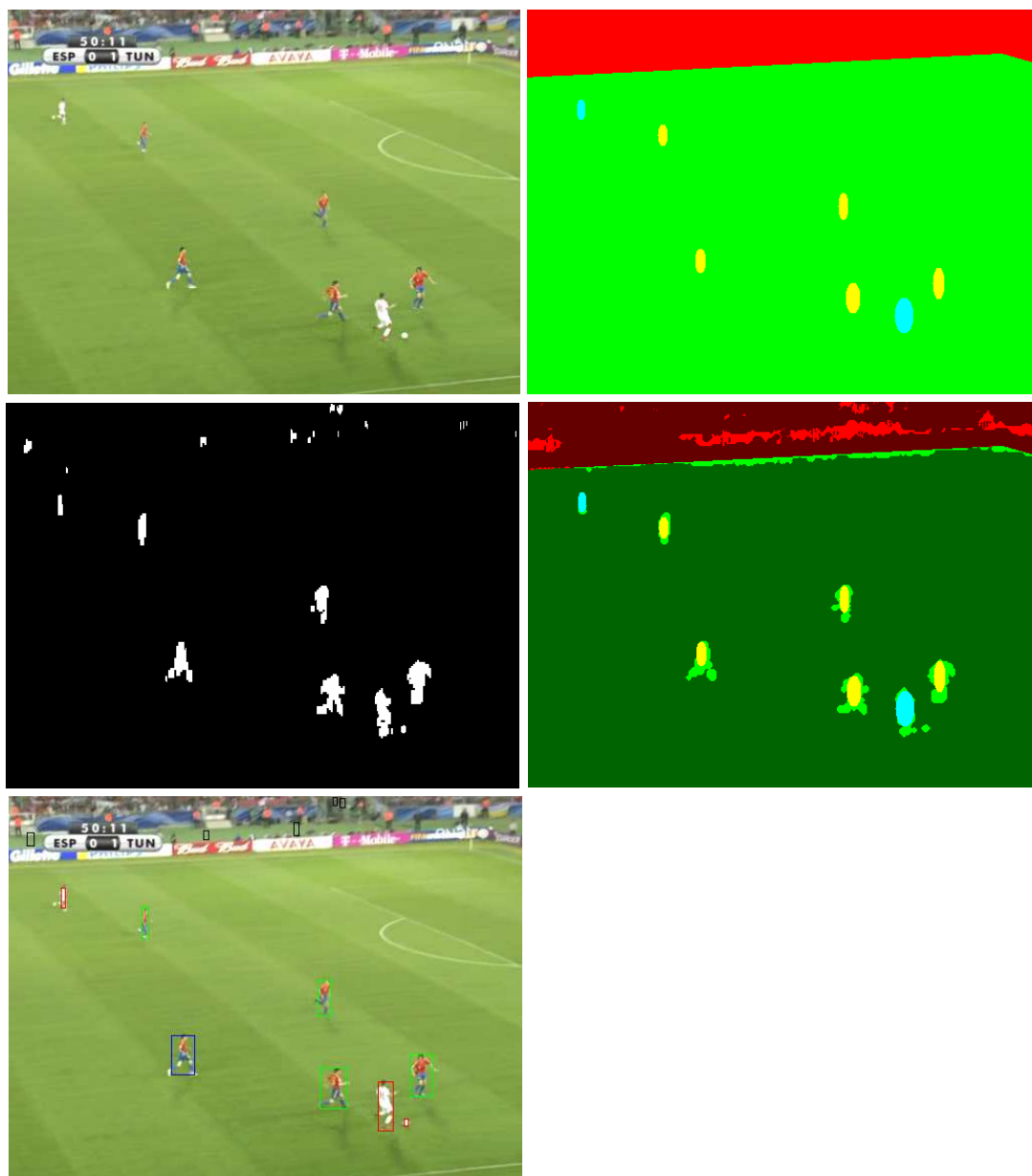


Figura B.5 – Exemplo

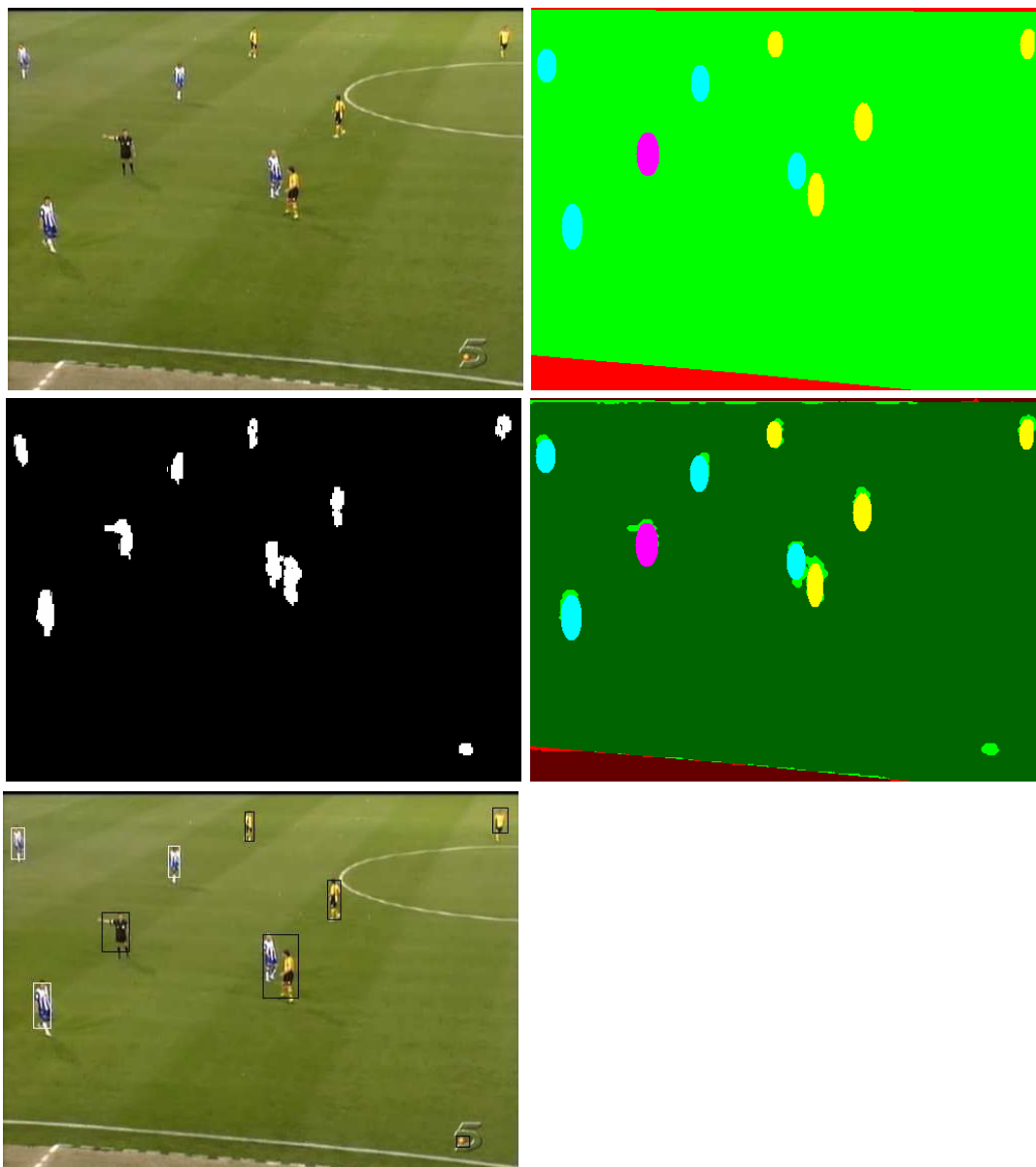


Figura B.6 – Exemplo

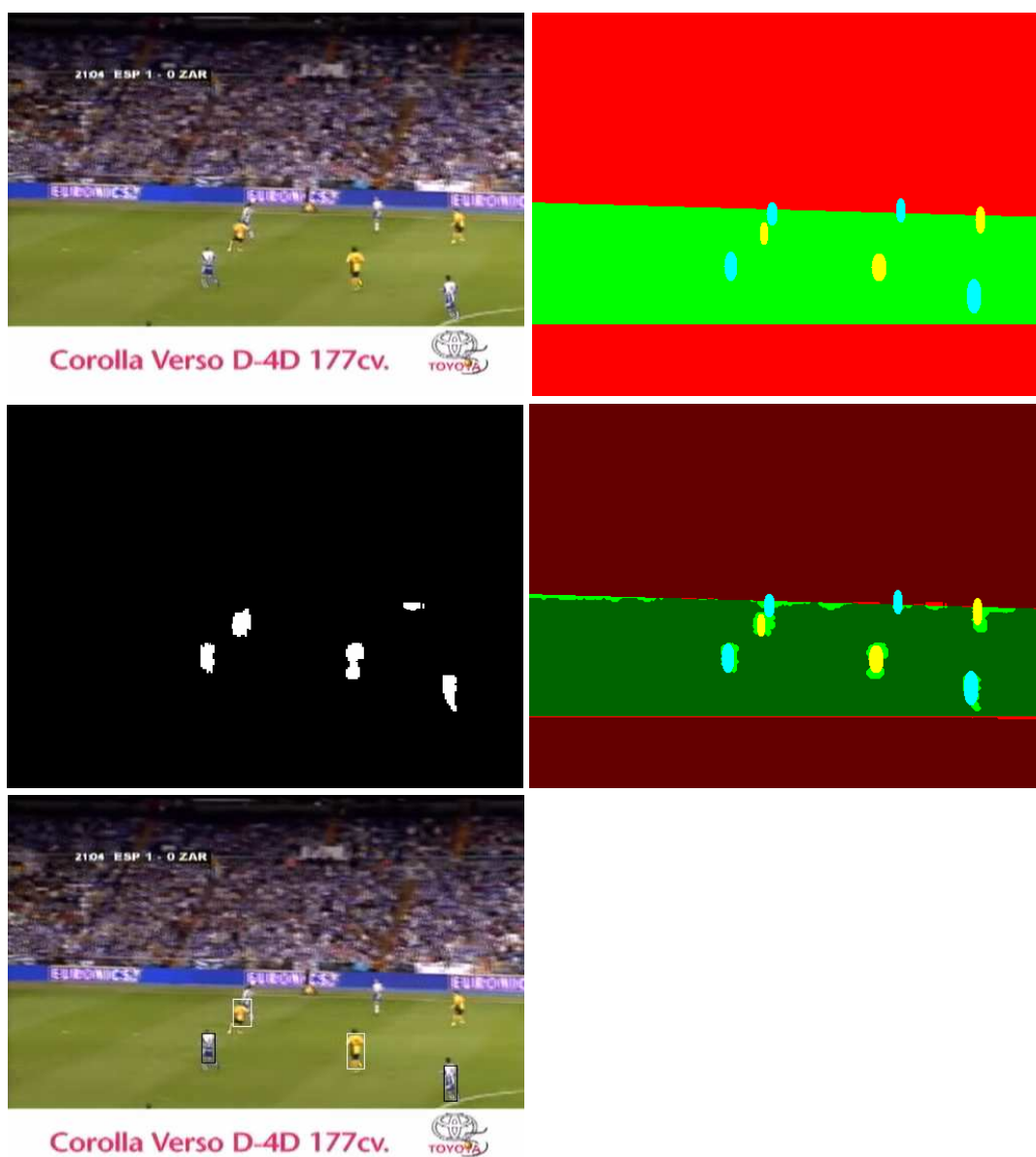


Figura B.7 – Exemplo

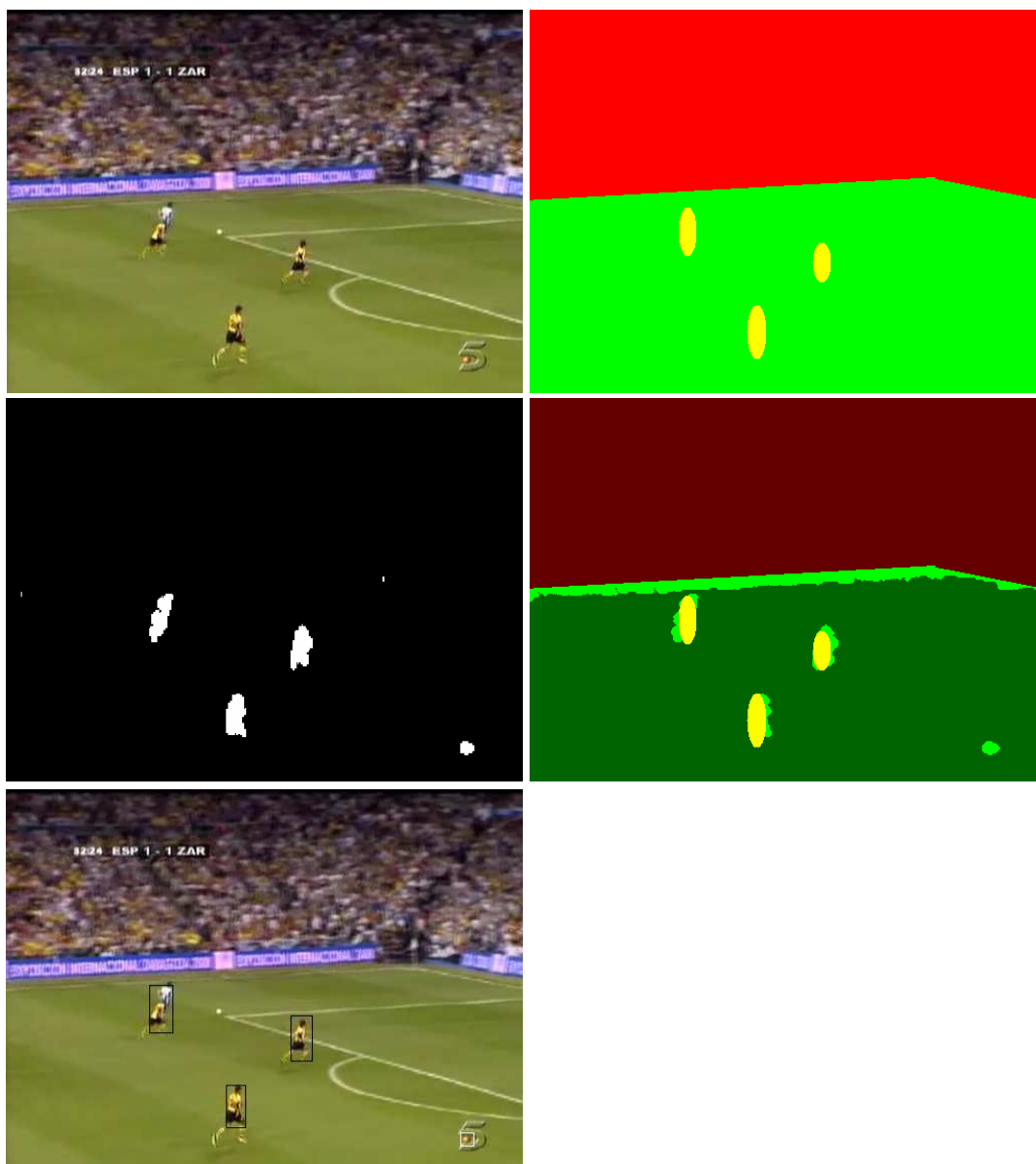


Figura B.8 – Exemplo

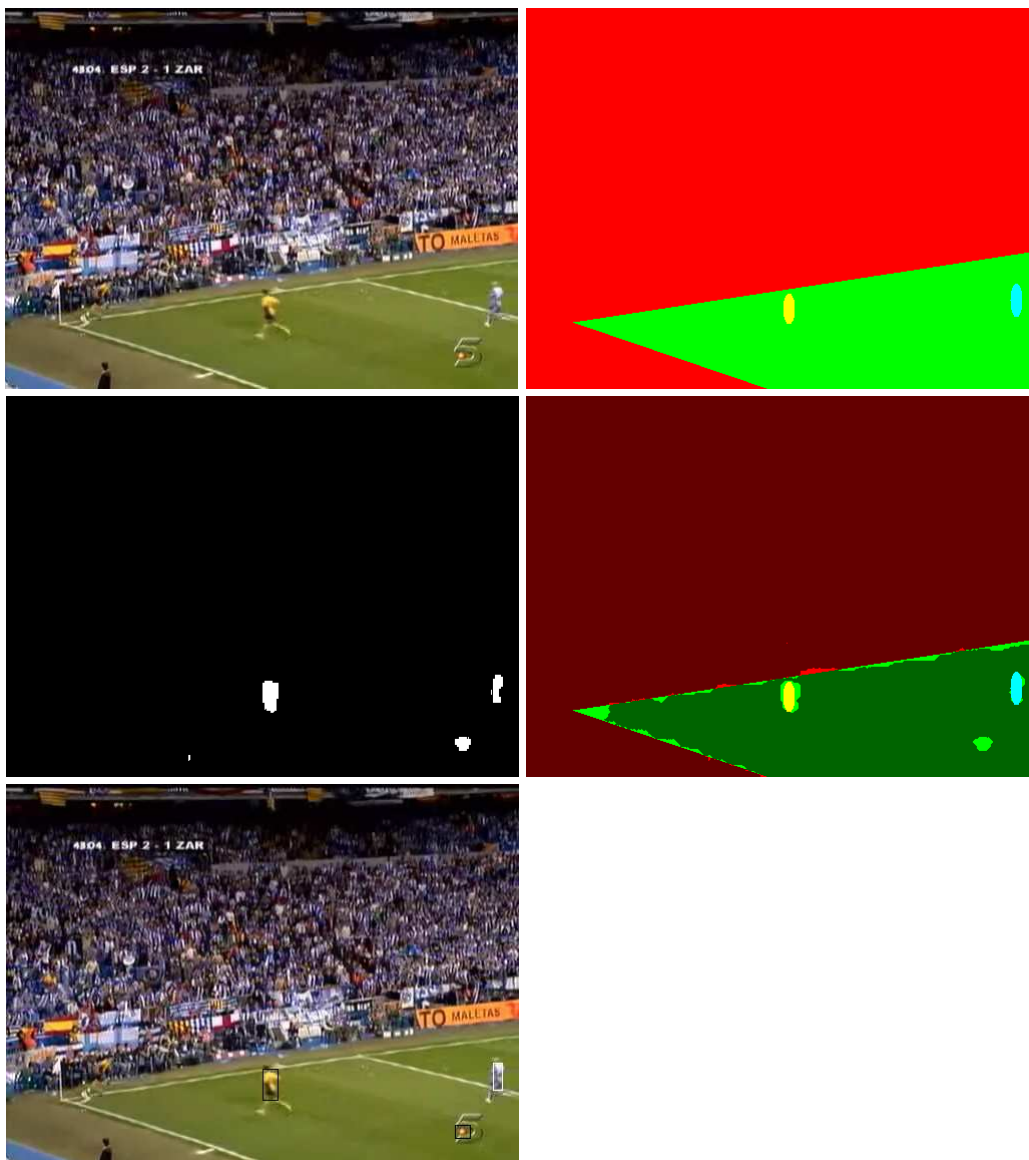


Figura B.9 – Exemplo

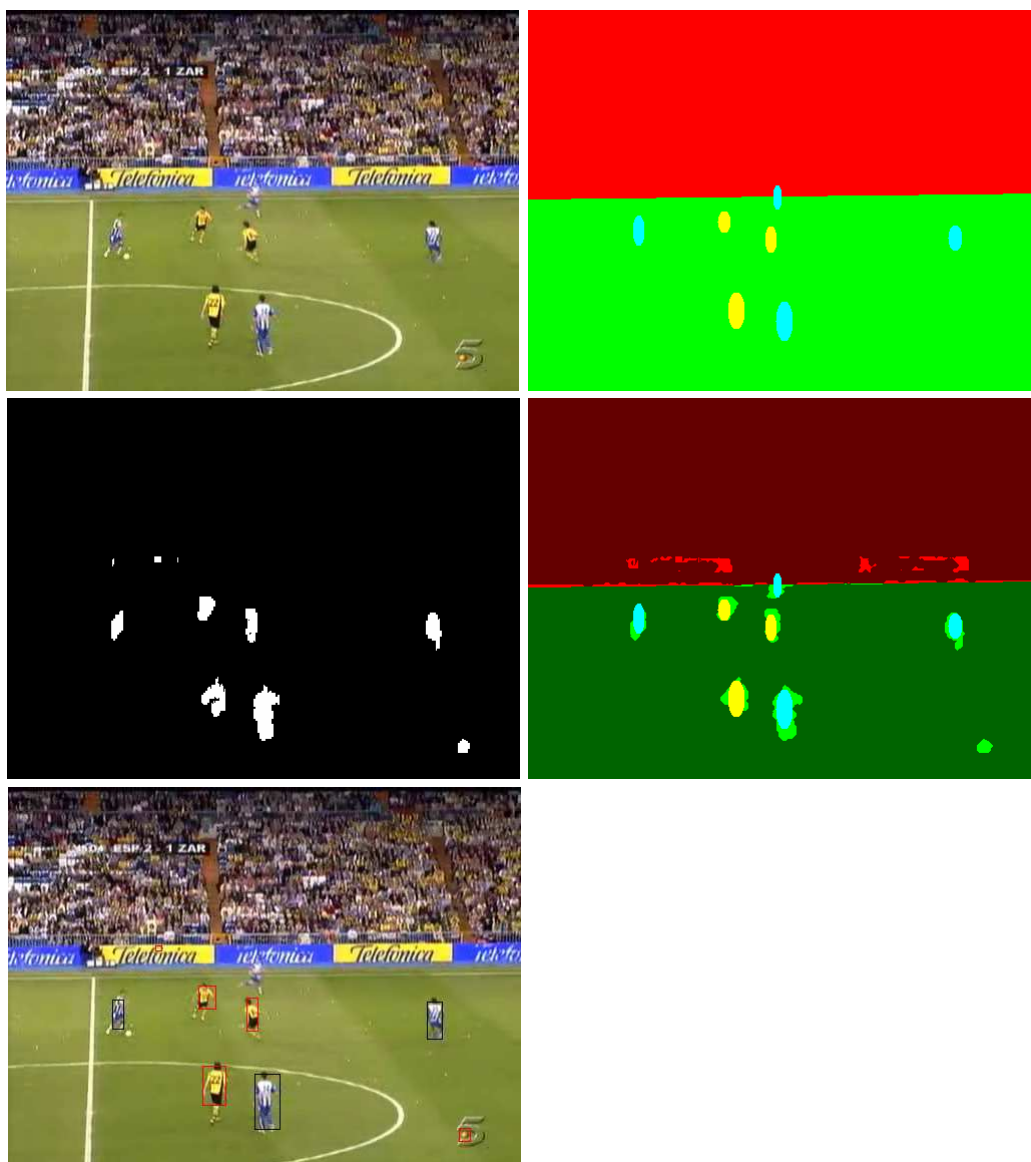


Figura B.10 – Exemplo

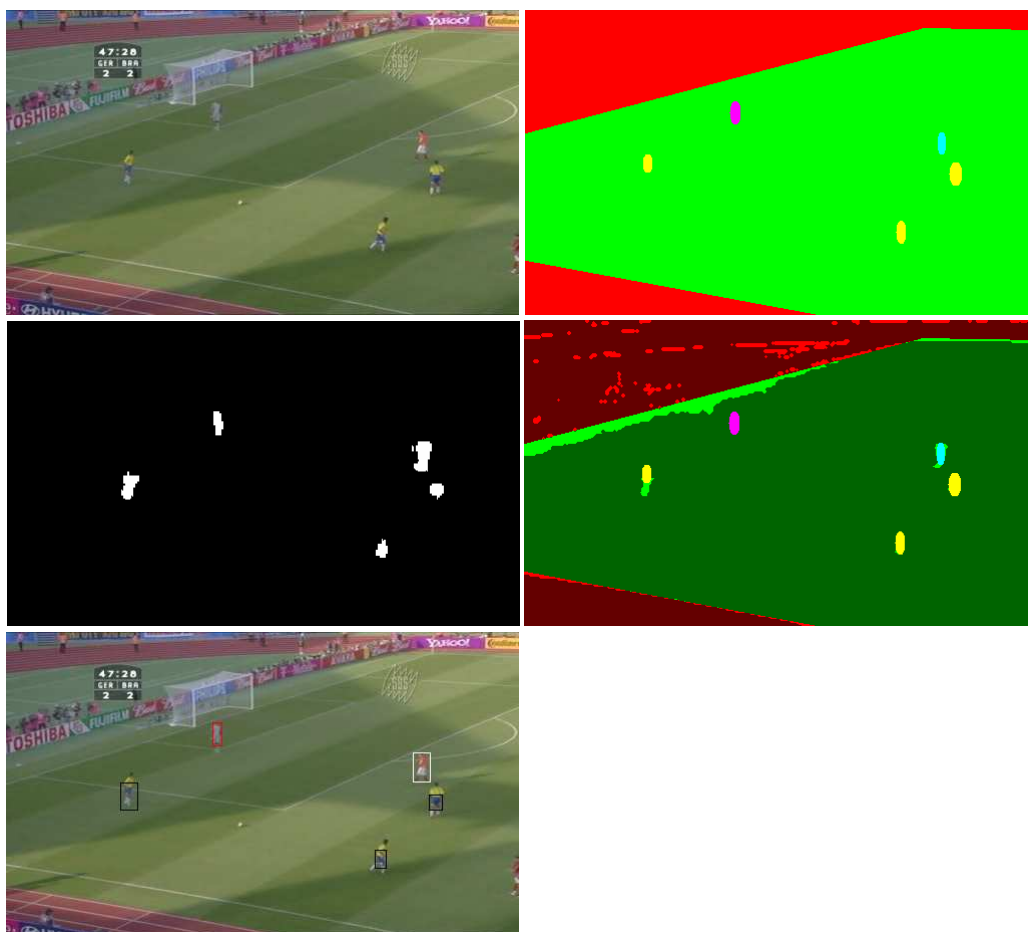


Figura B.11 – Exemplo

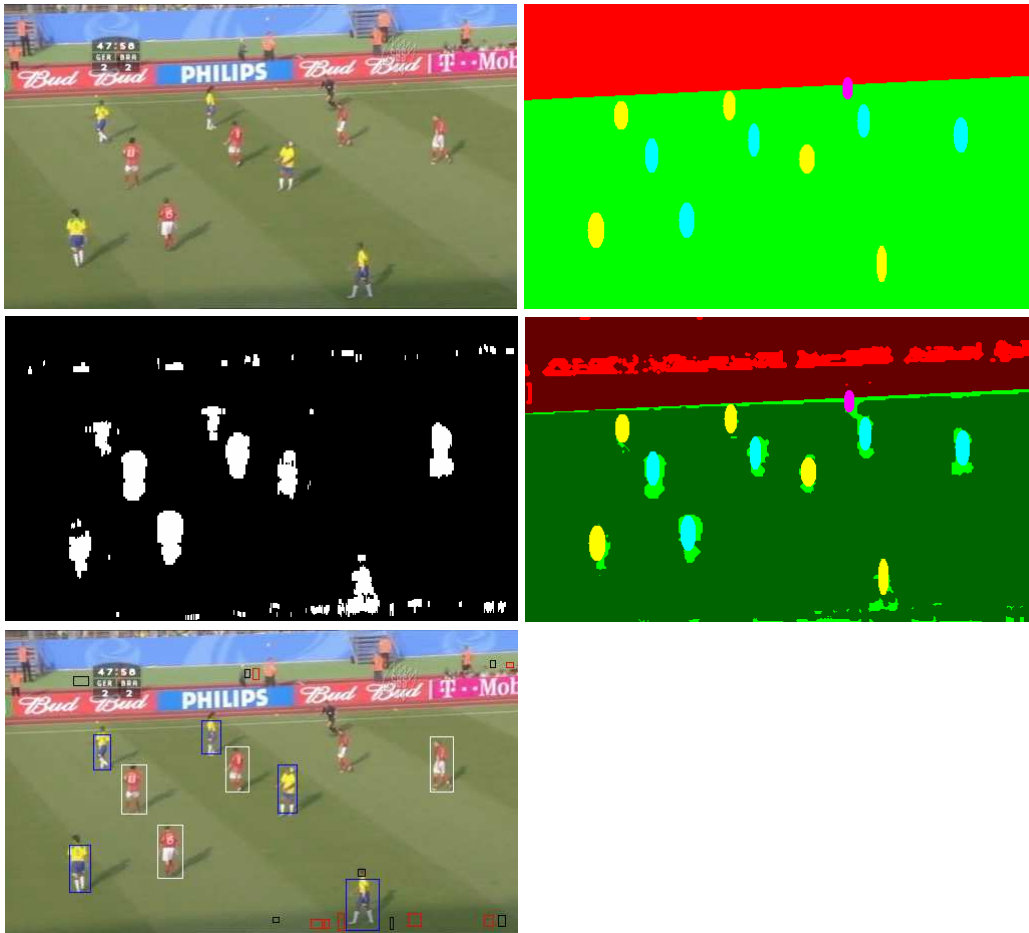


Figura B.12 – Exemplo

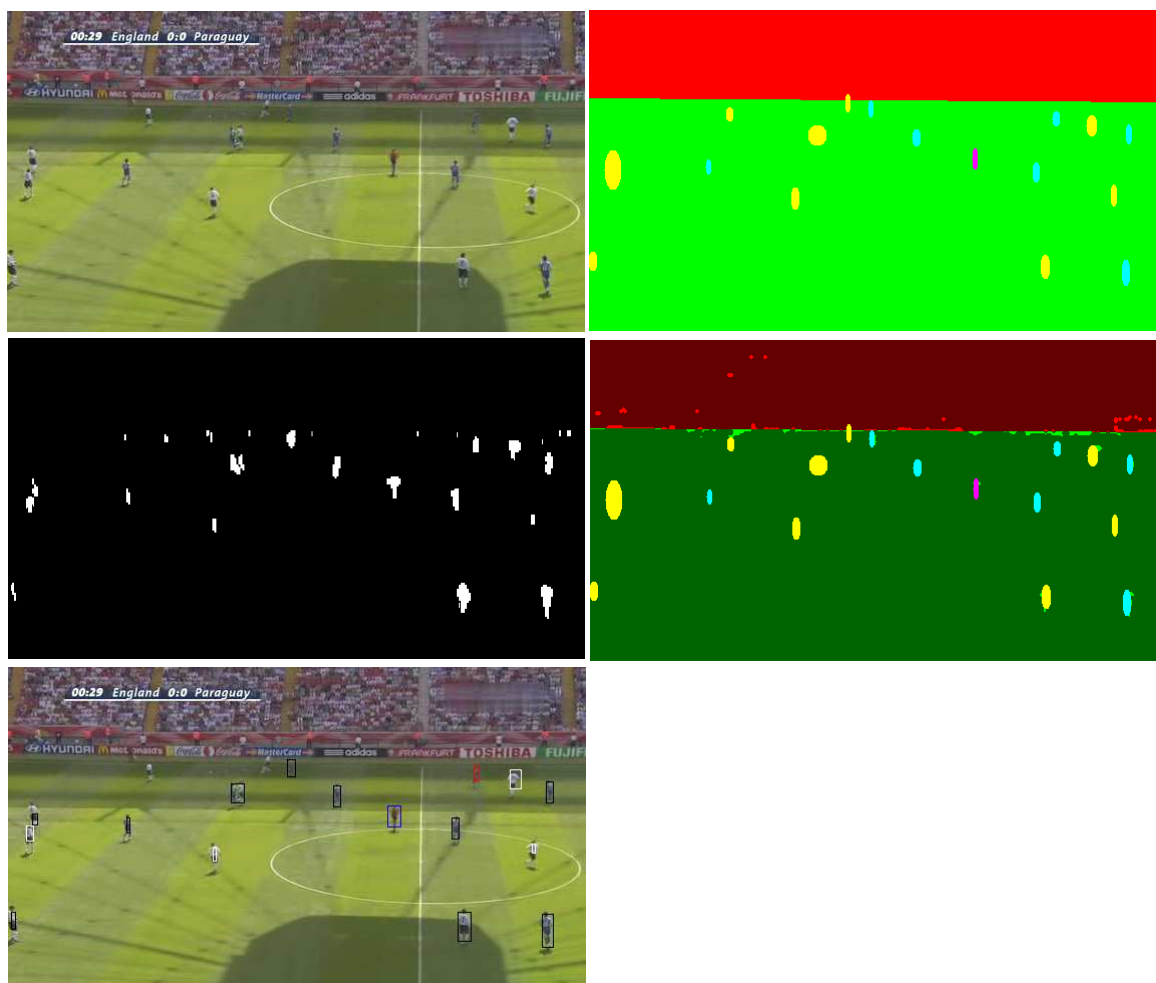


Figura B.13 – Exemplo

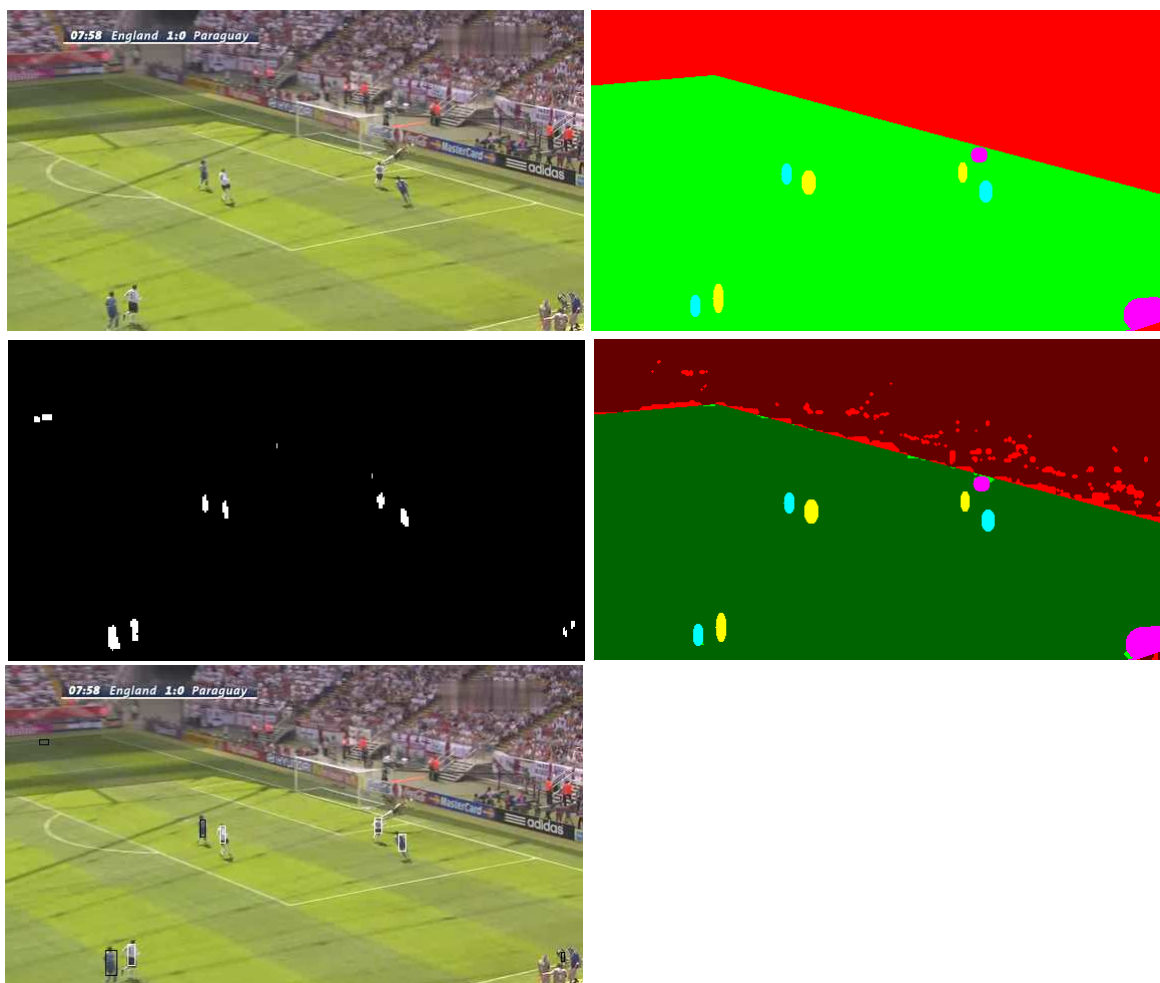


Figura B.14 – Exemplo

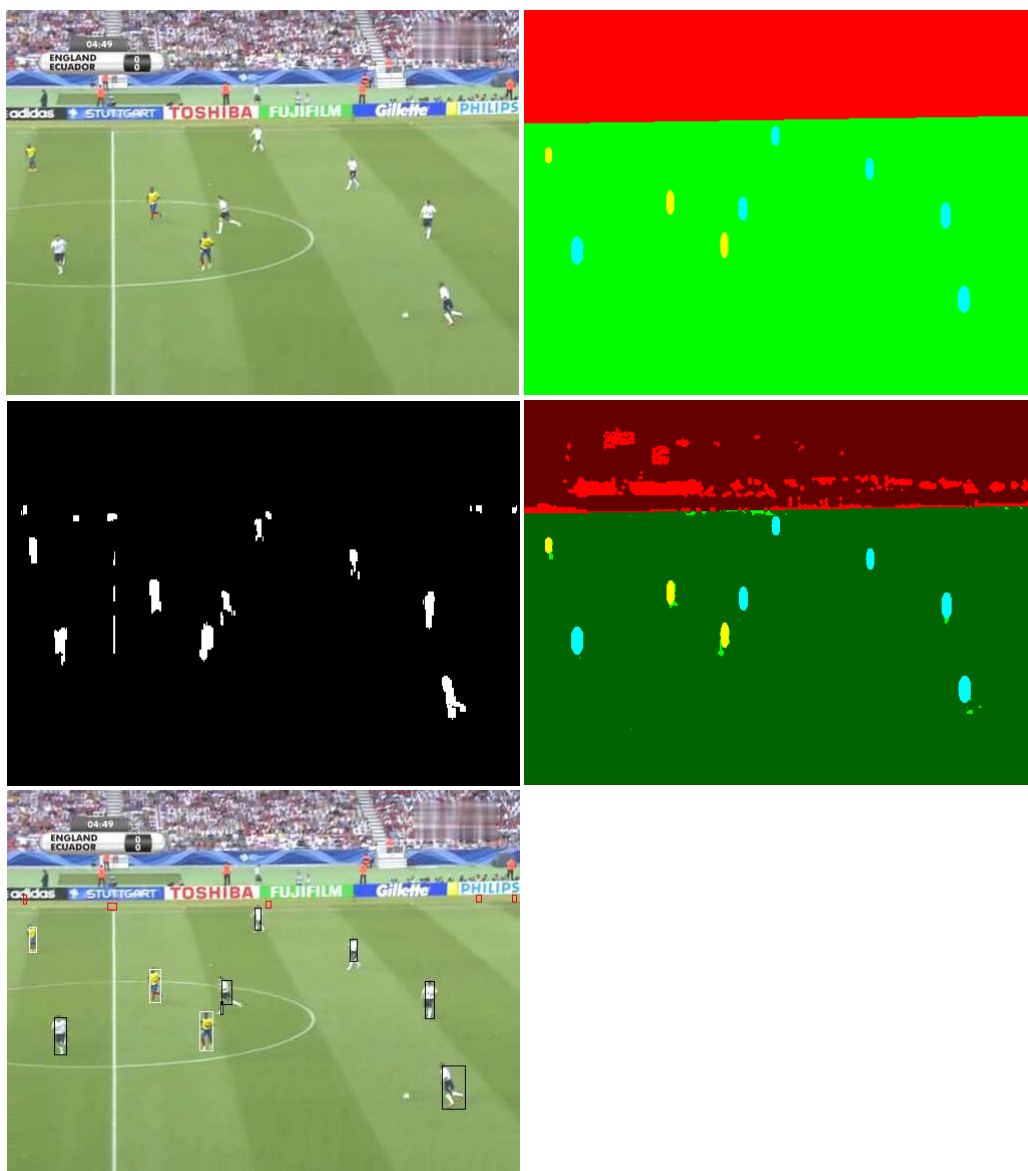


Figura B.15 – Exemplo

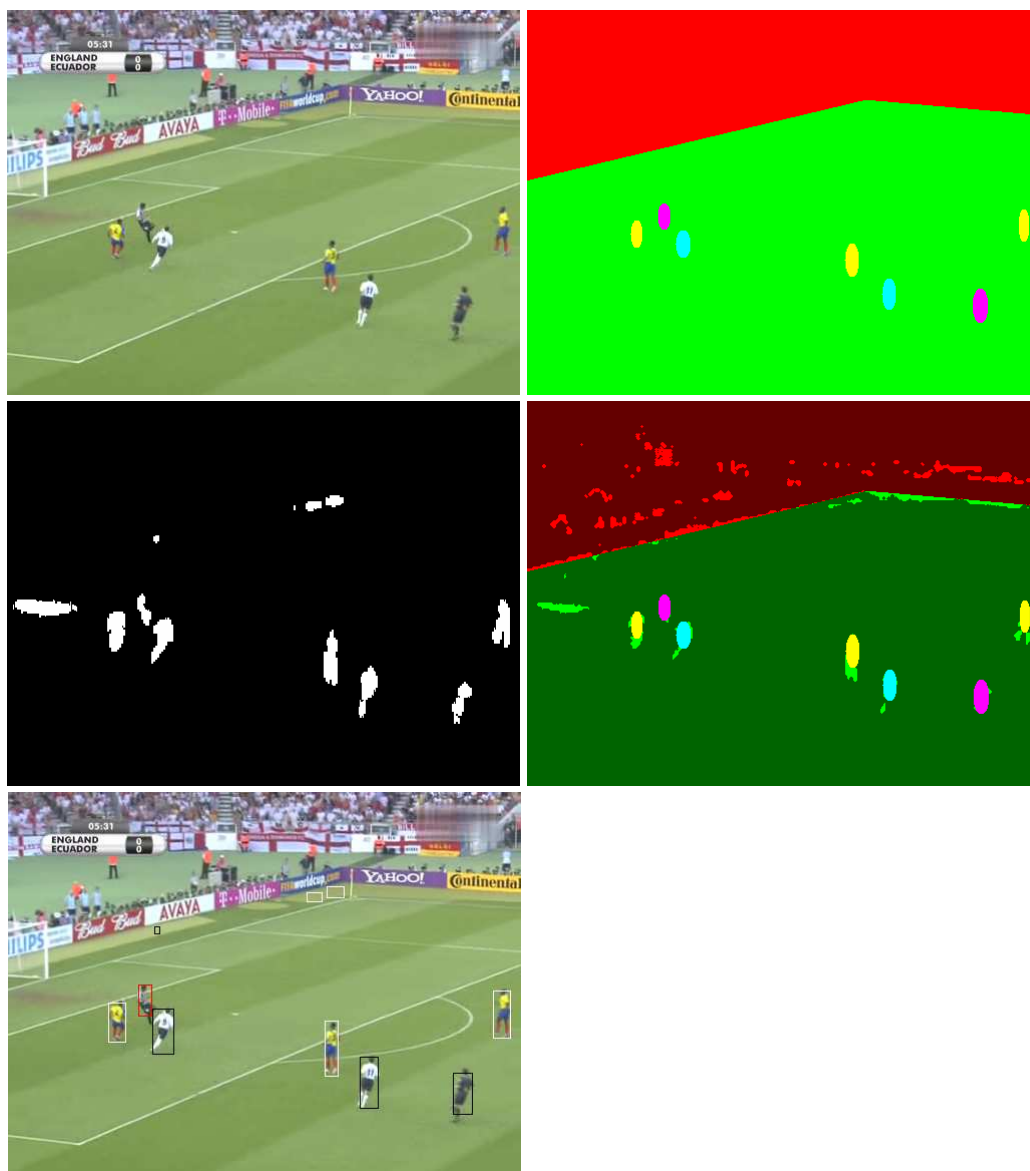


Figura B.16 – Exemplo

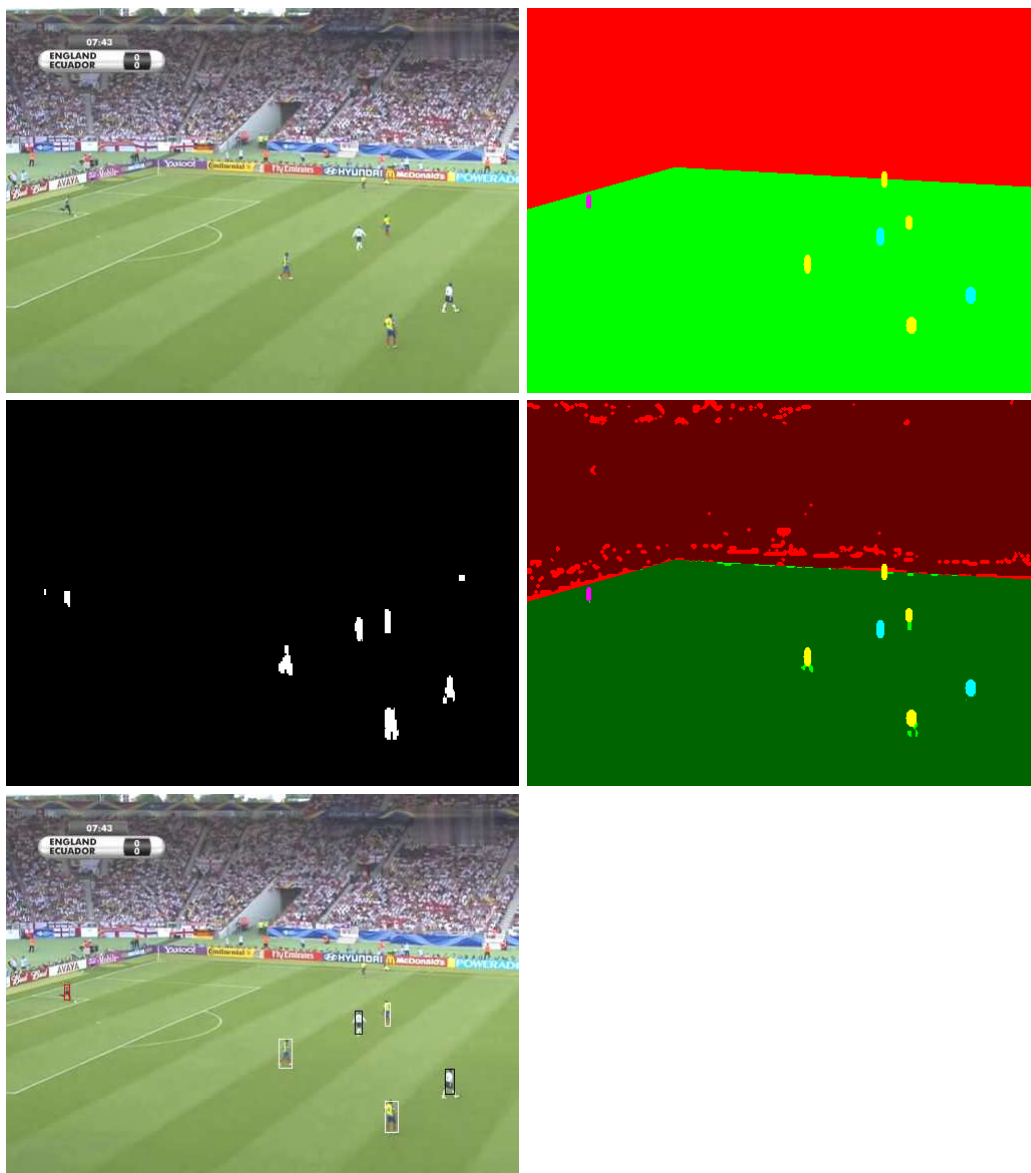


Figura B.17 – Exemplo

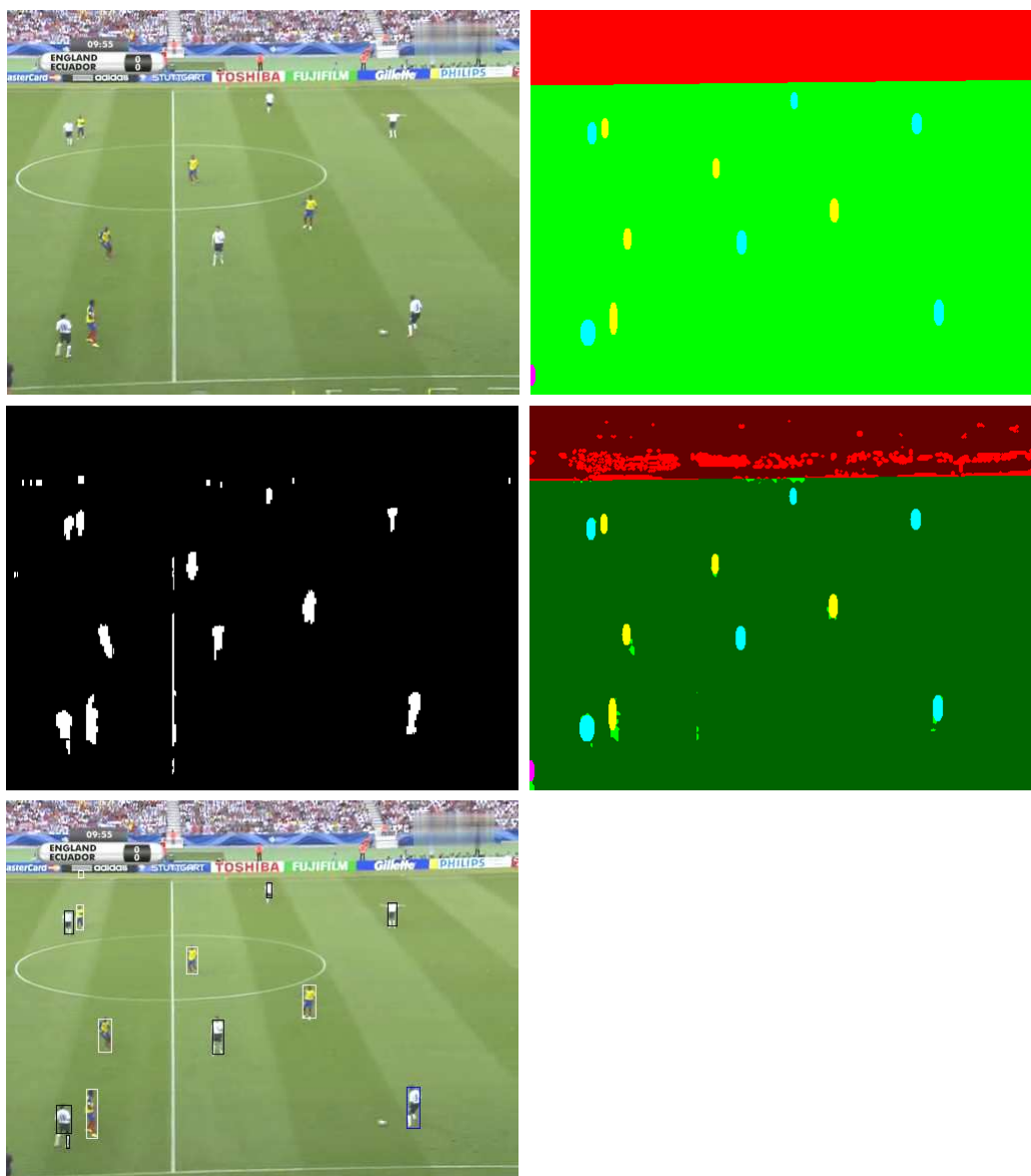


Figura B.18 – Exemplo

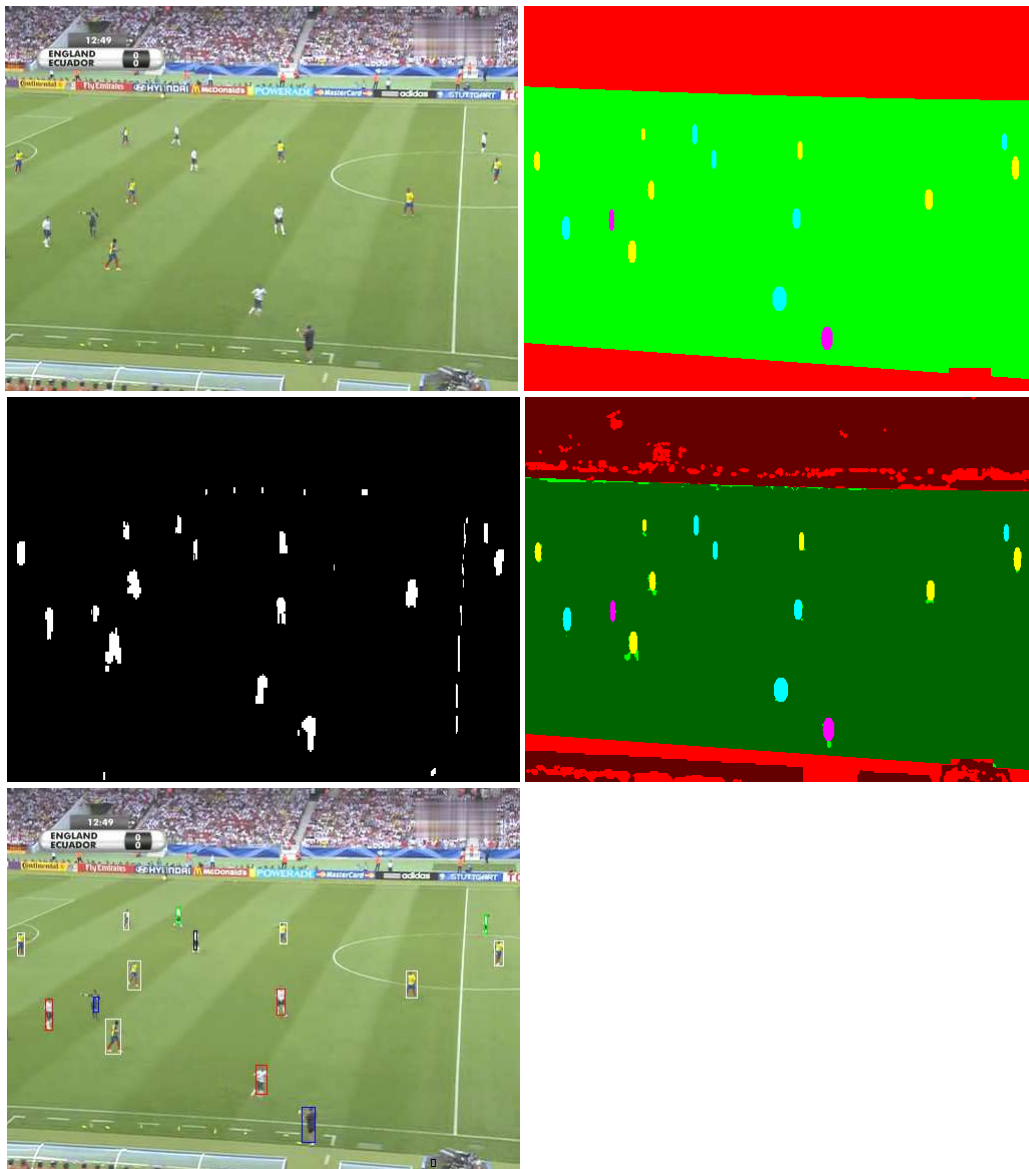


Figura B.19 – Exemplo

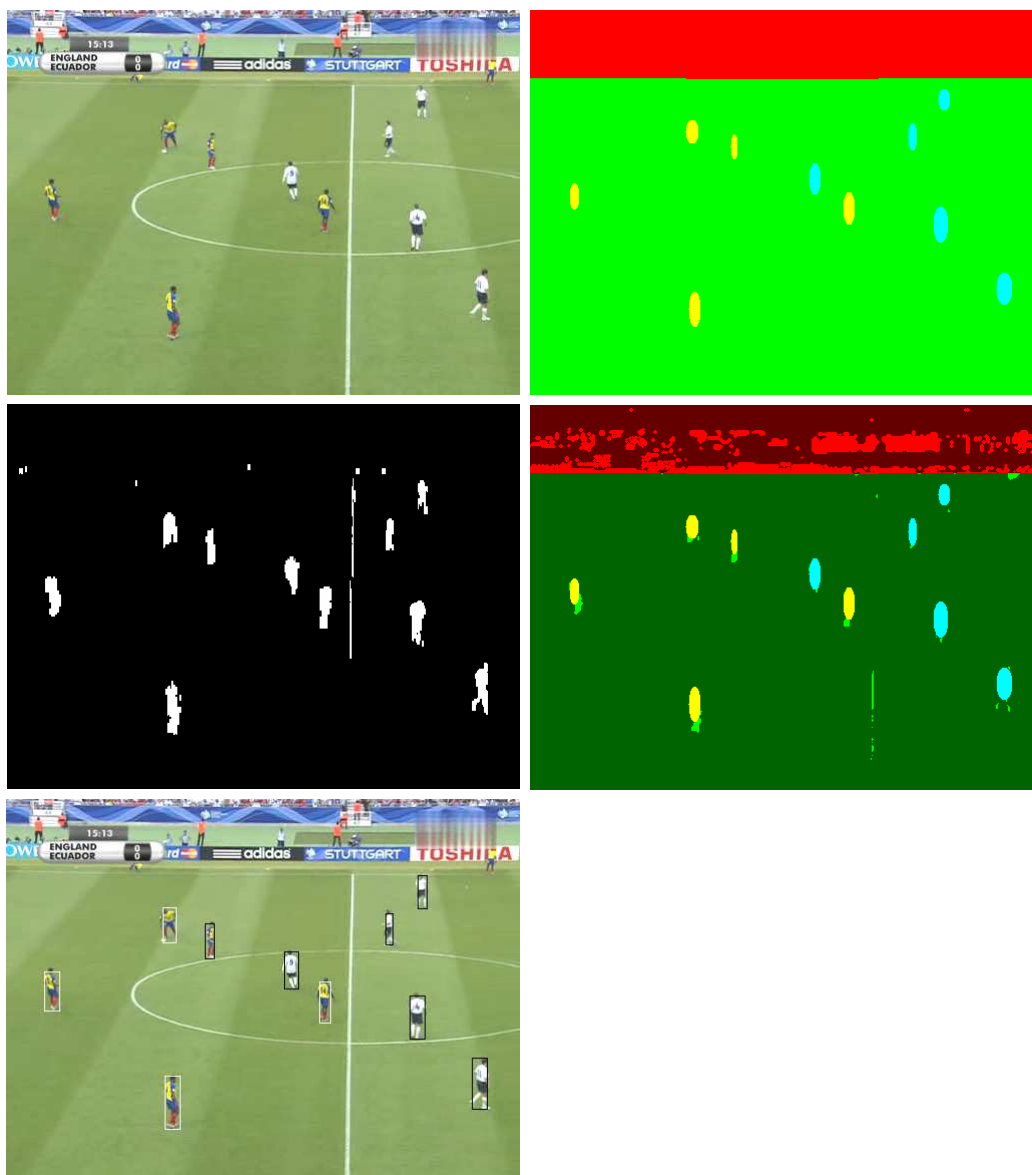


Figura B.20 – Exemplo

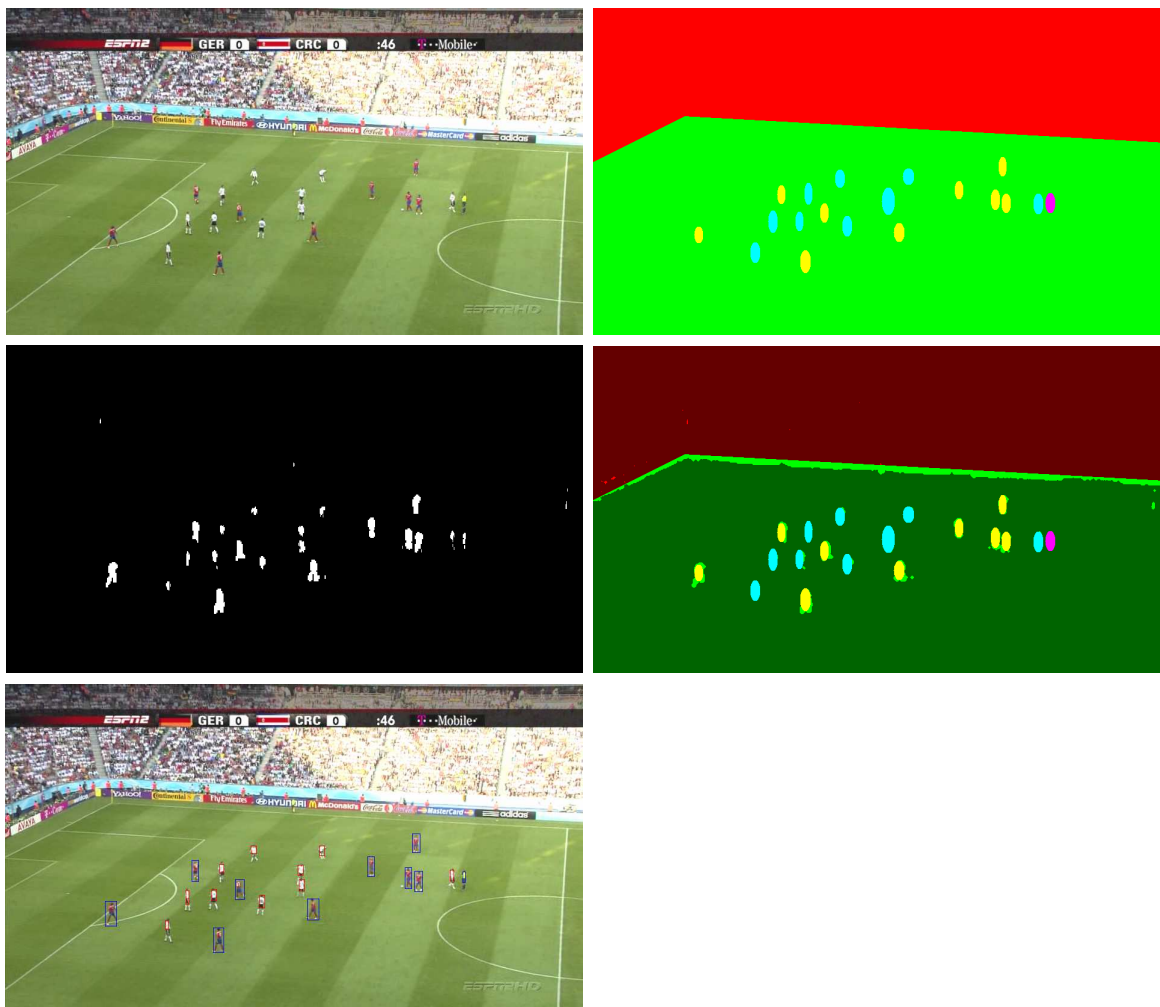


Figura B.21 – Exemplo

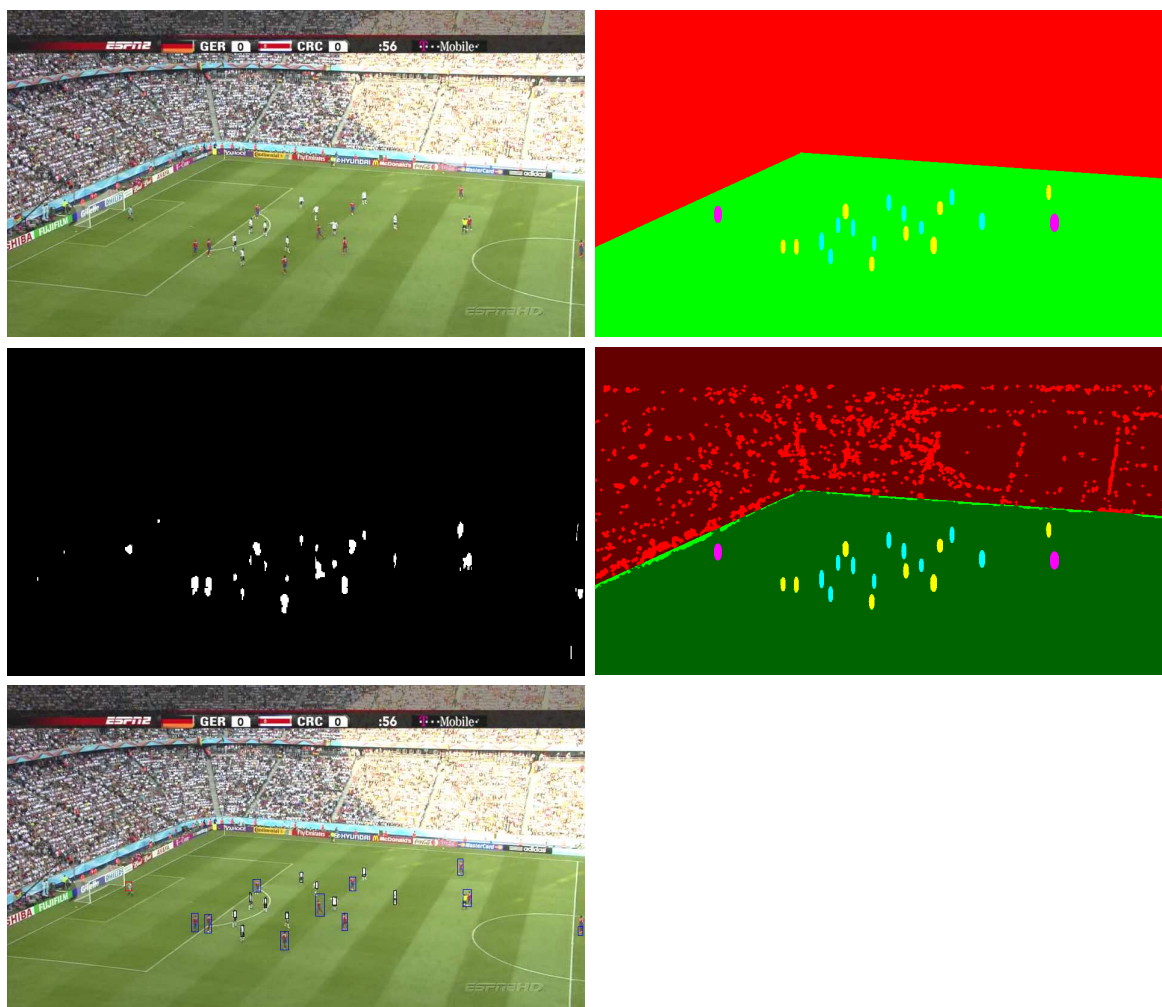


Figura B.22 – Exemplo

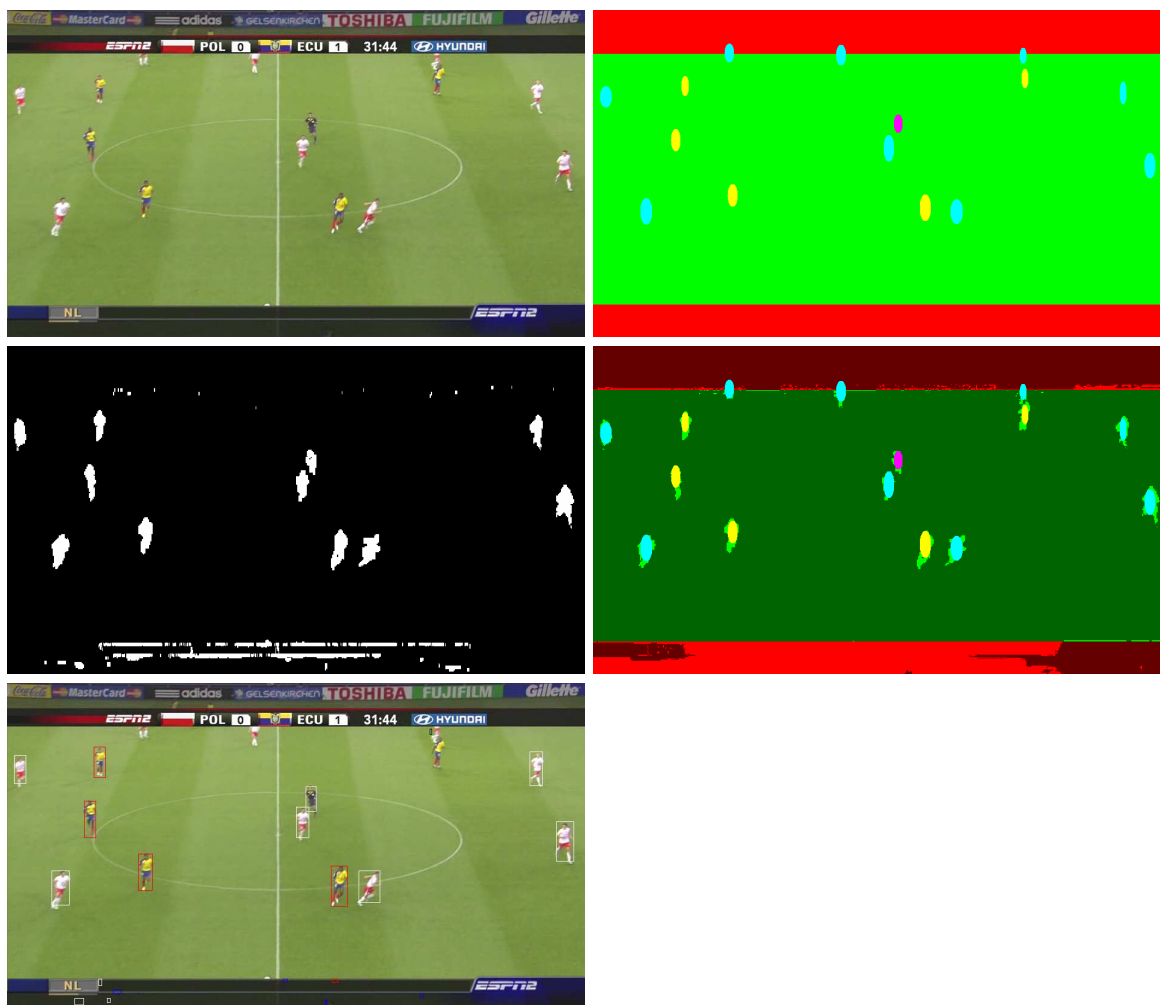


Figura B.23 – Exemplo

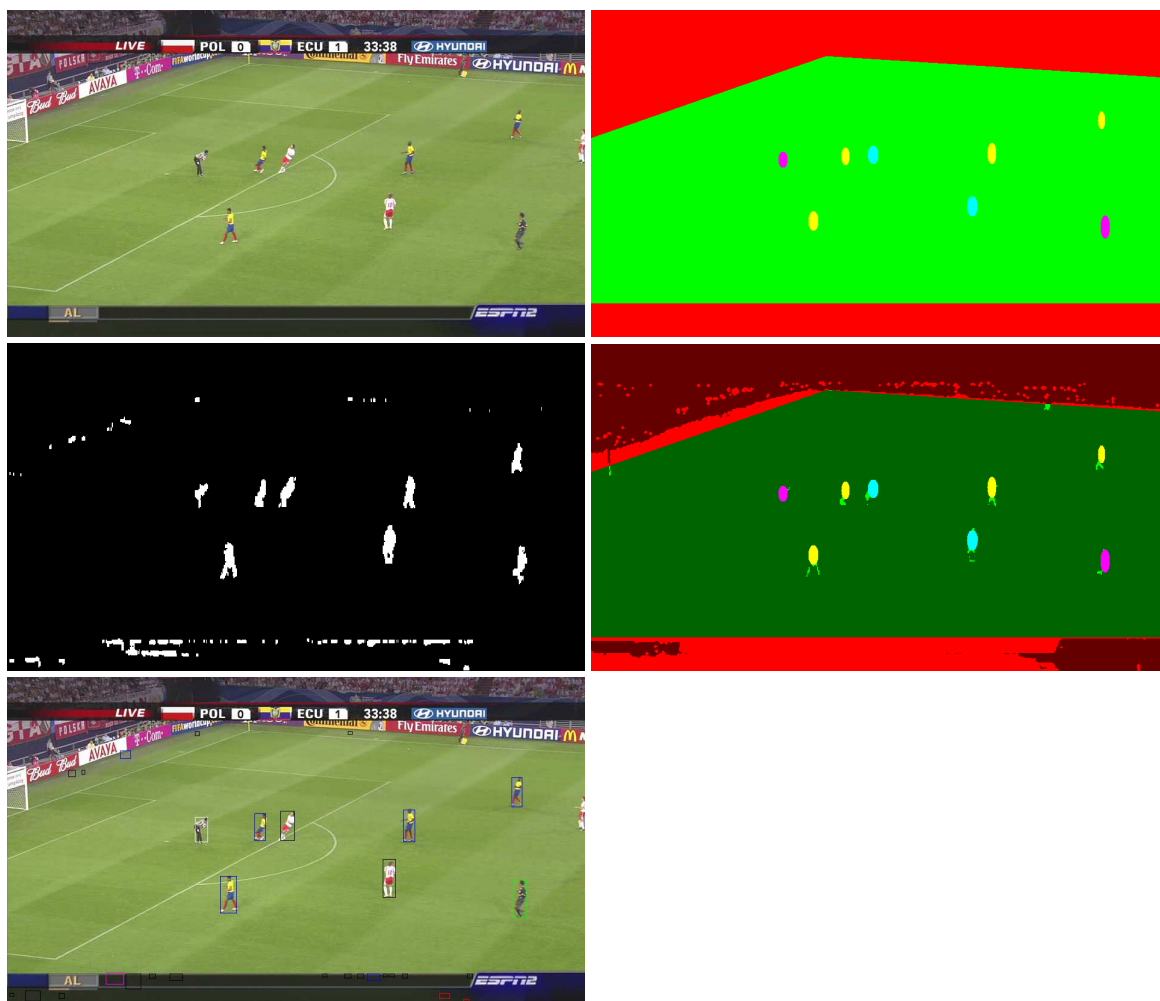


Figura B.24 – Exemplo

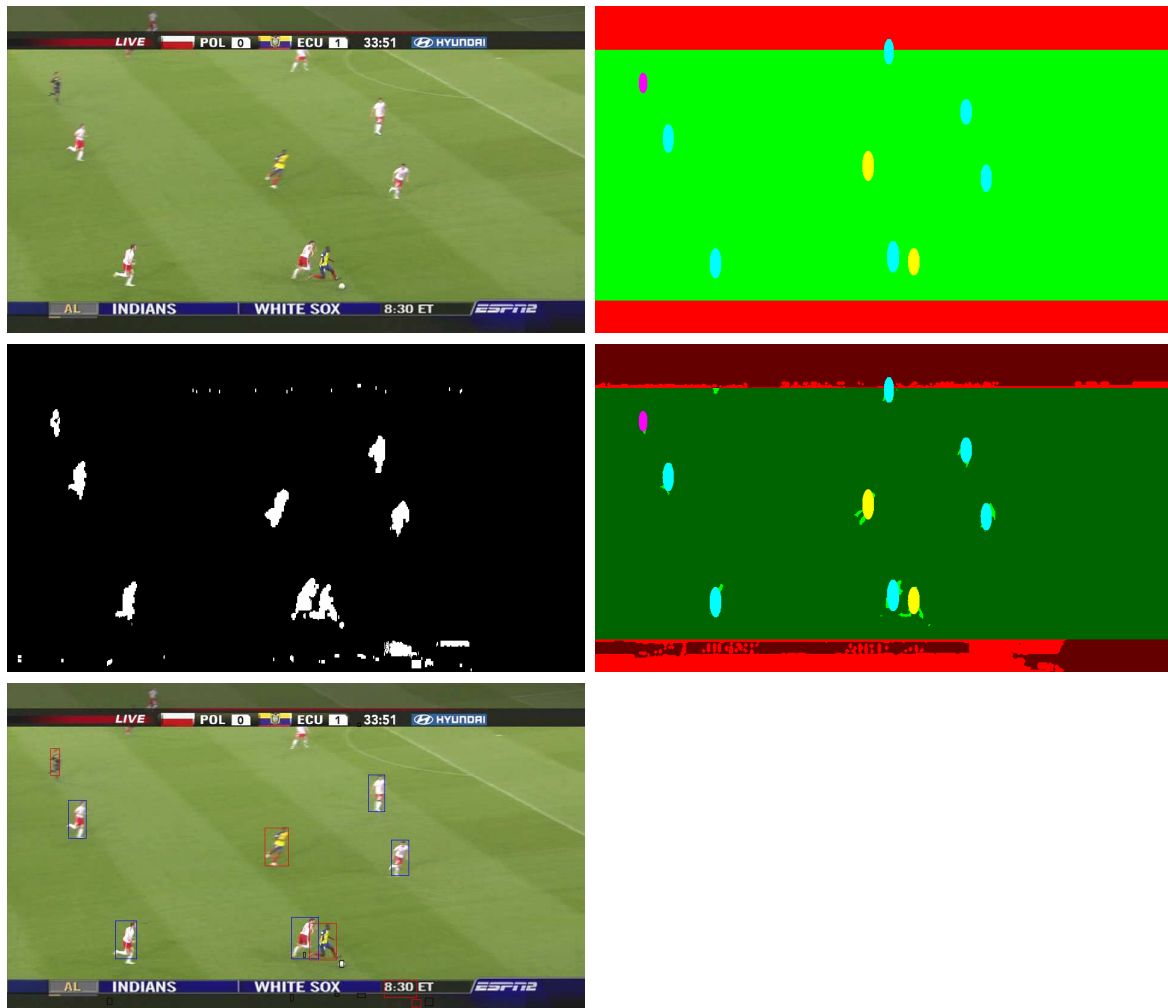


Figura B.25 – Exemplo

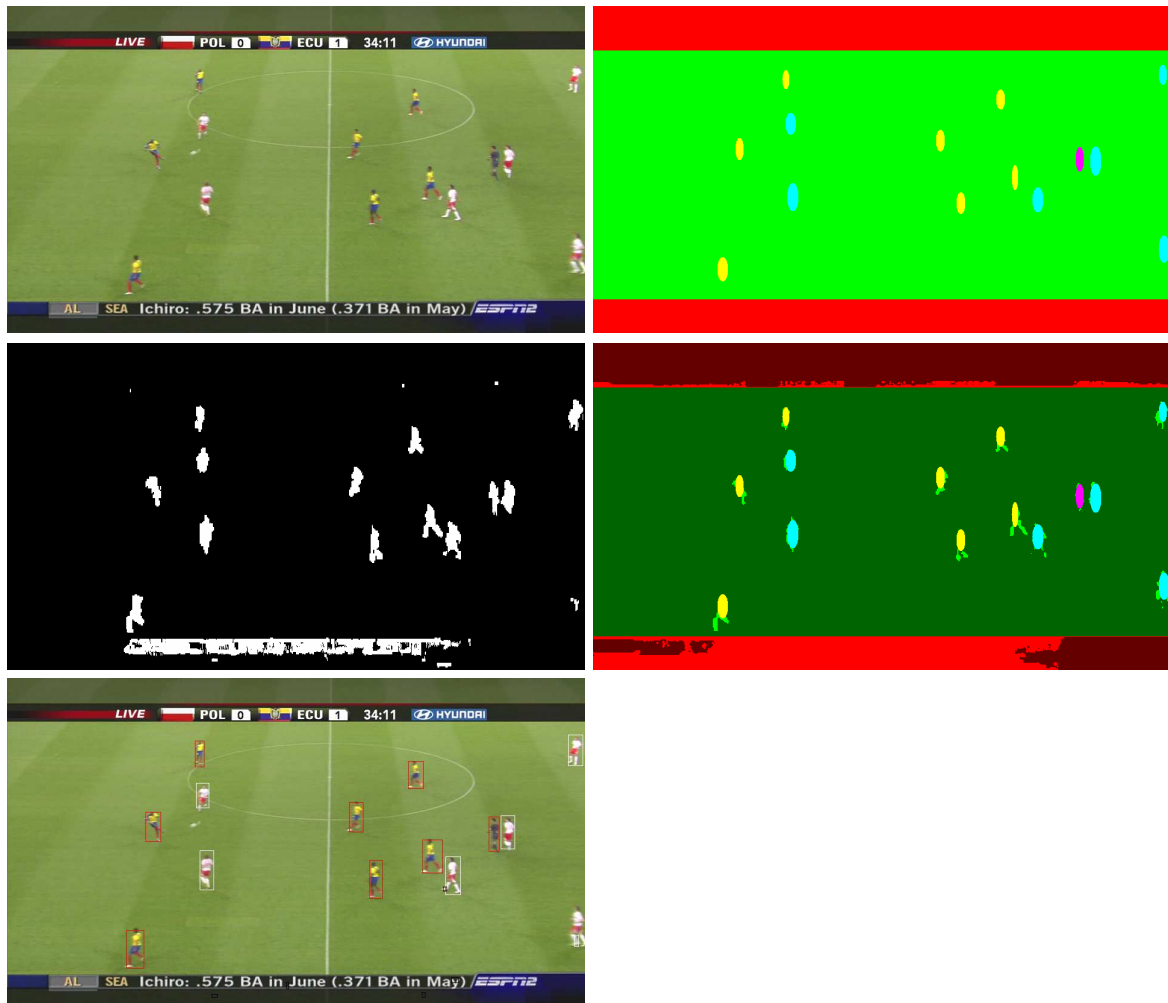


Figura B.26 – Exemplo

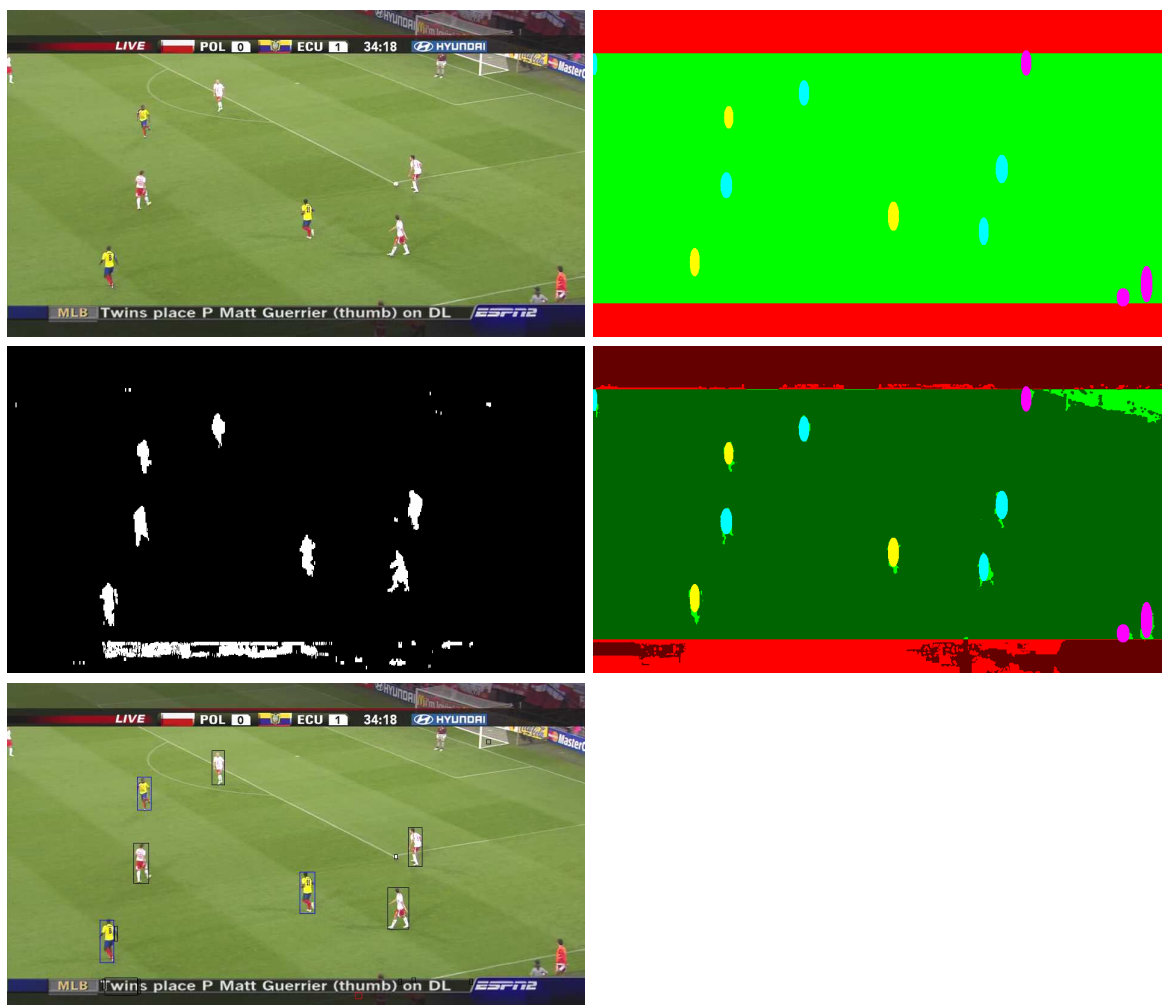


Figura B.27 – Exemplo

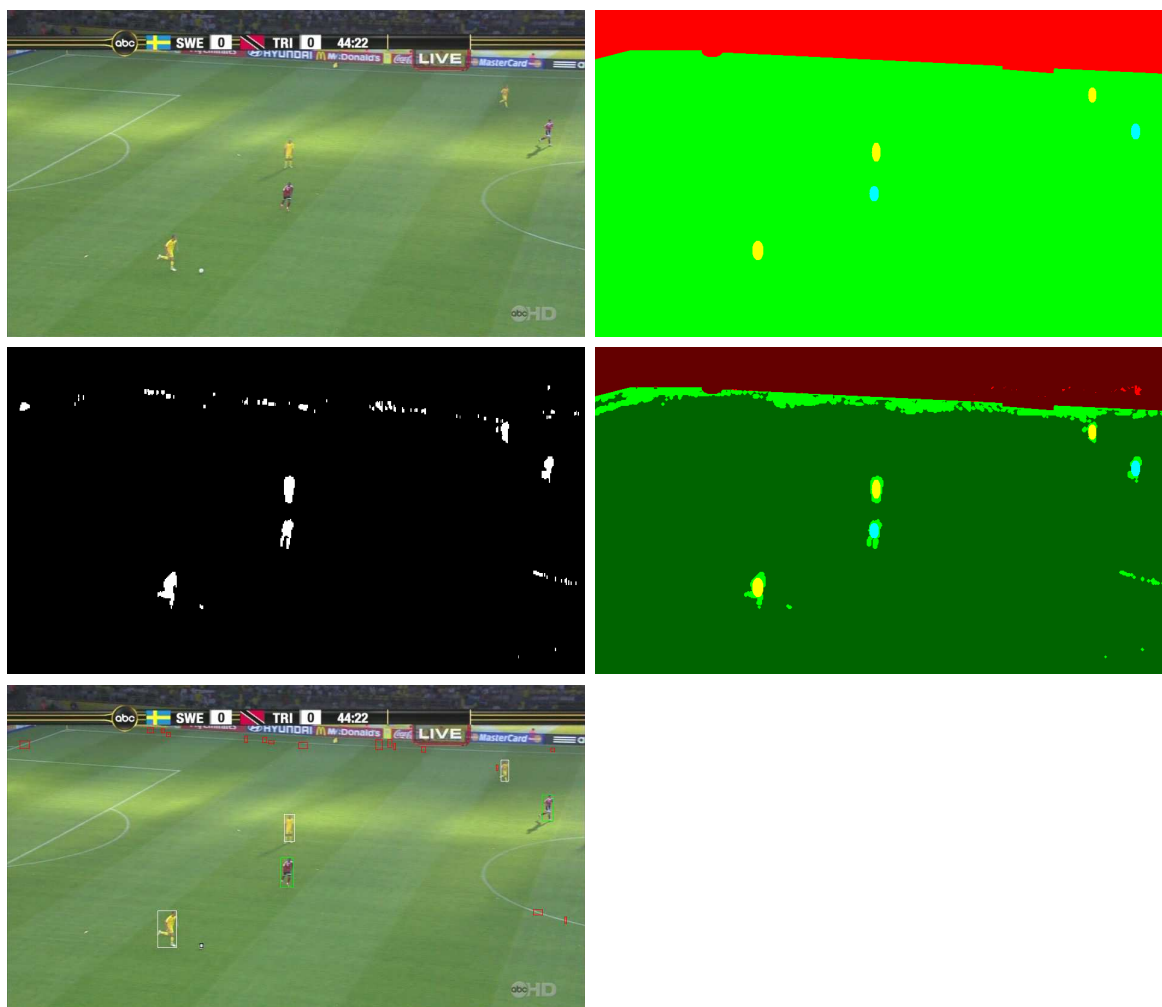


Figura B.28 – Exemplo

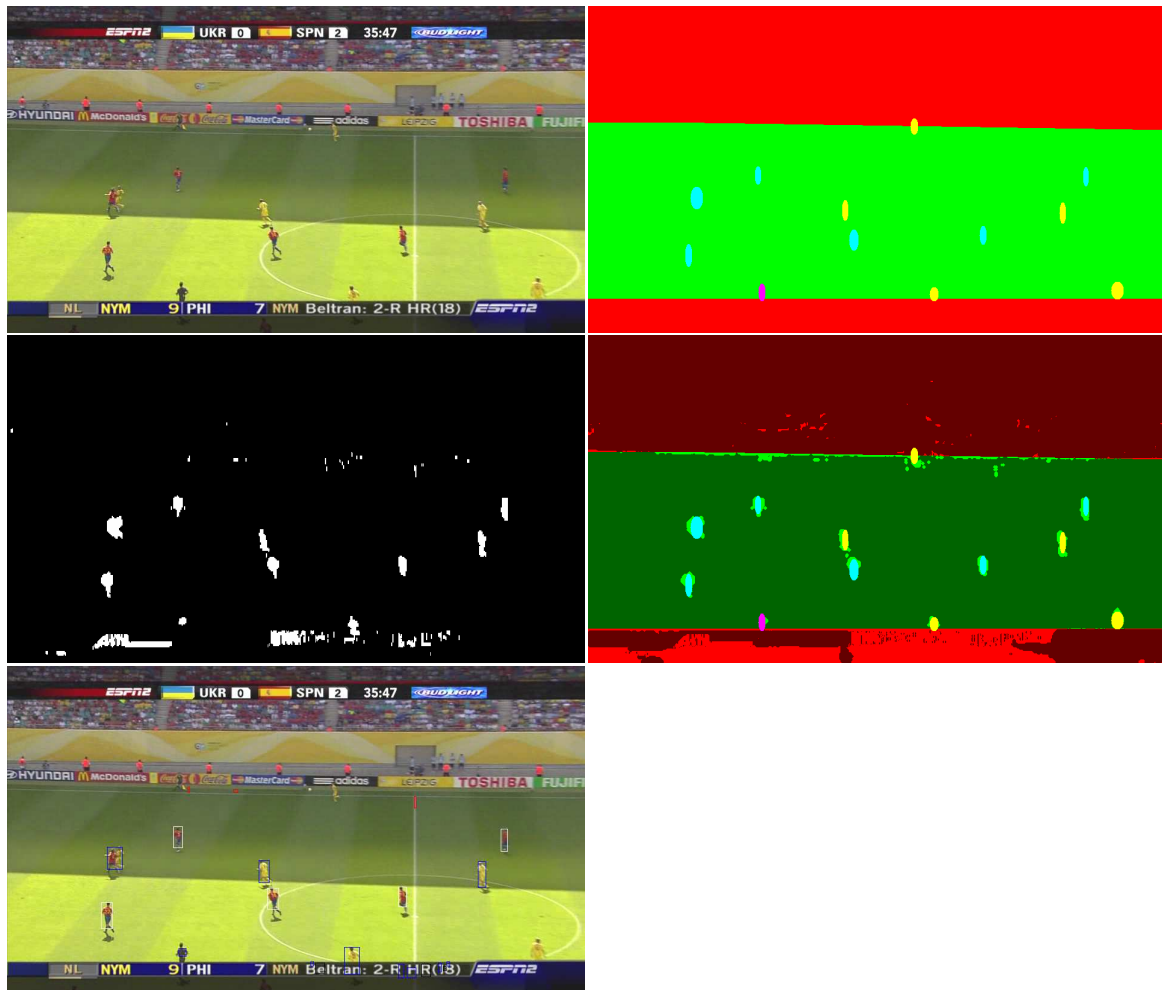


Figura B.29 – Exemplo

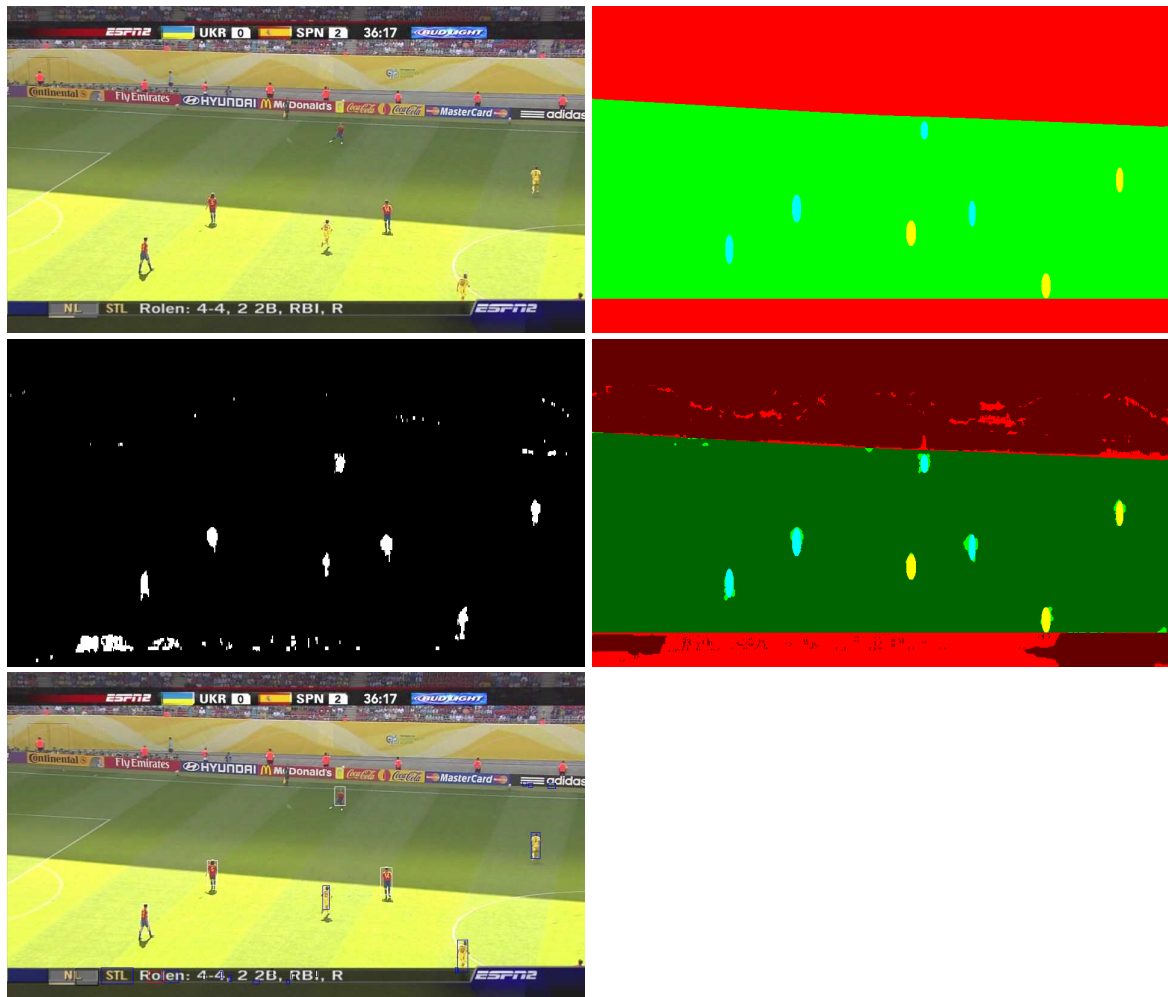


Figura B.30 – Exemplo

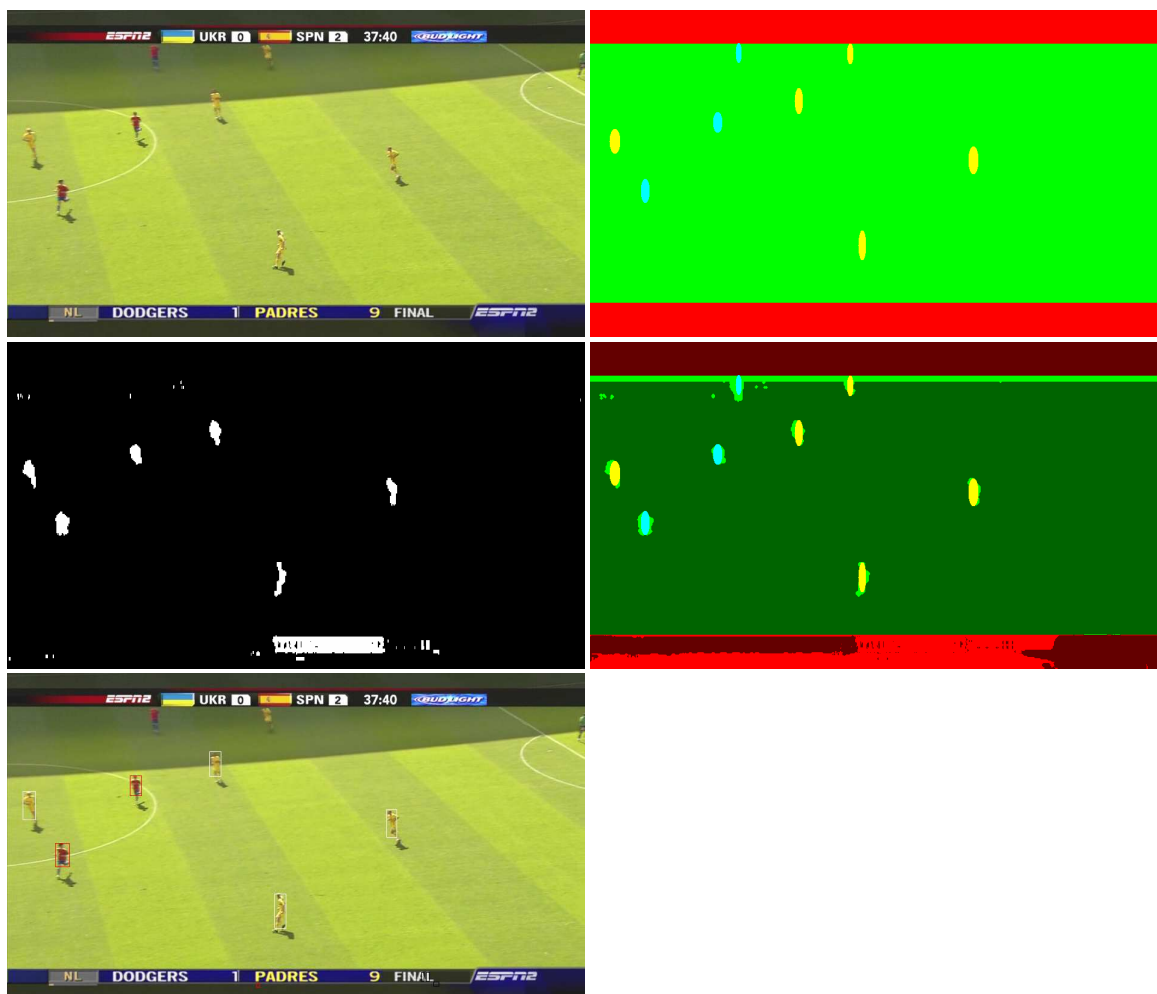


Figura B.31 – Exemplo

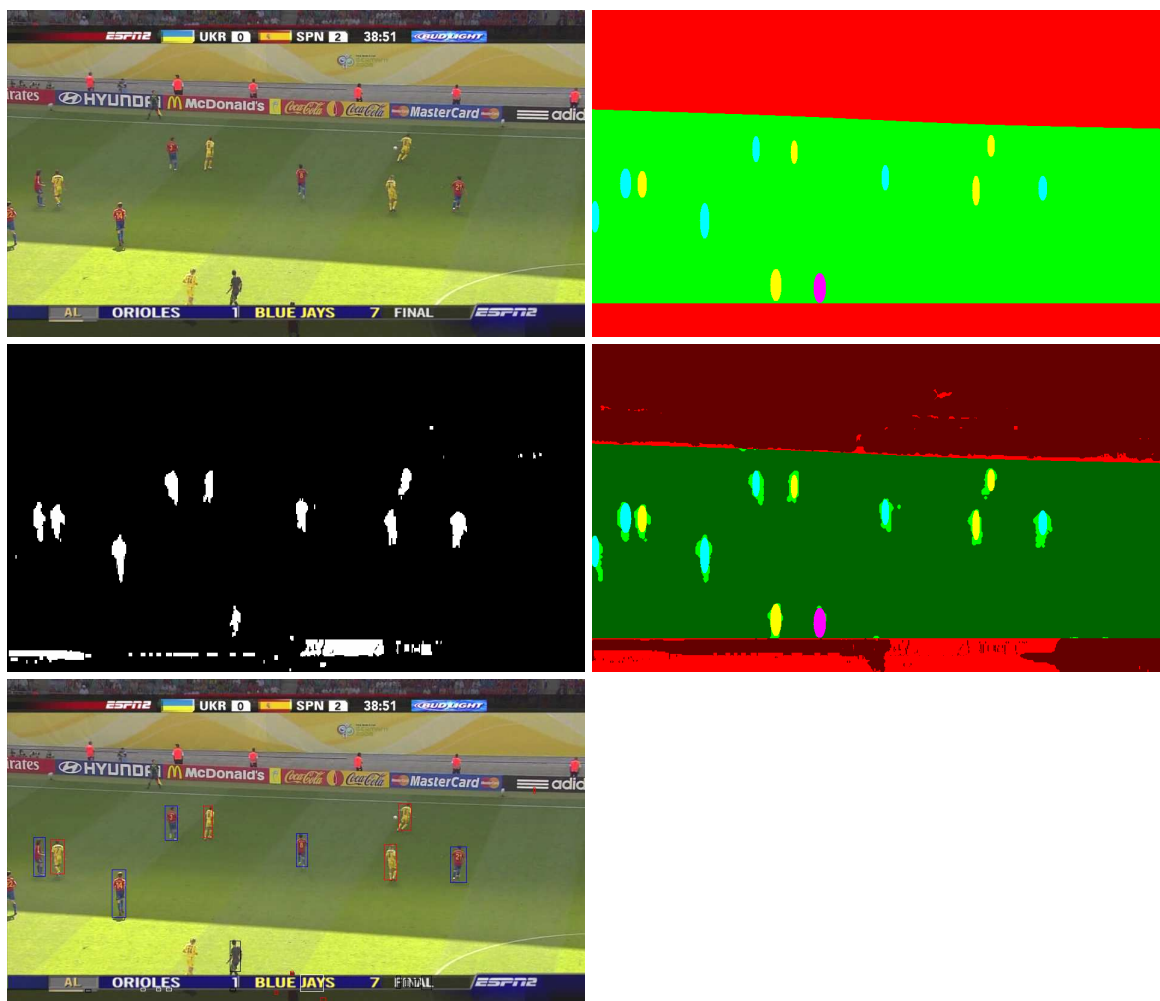


Figura B.32 – Exemplo

Apêndice C

Índices de qualidade para novos pesos

Característica	Peso
Campo	35%
Adjacencia	35%
JSC	10%
JCC	10%
RCCP	10%

Tabela C.1 – Tabela de pesos (pesos originais)

Vídeo	CMPO	ADJ	JSC1	JSC2	JCC1	JCC2	RCCP	IQ1	IQ2
1	98,32%	80,75%	76,92%	82,47%	58,62%	63,75%	64,06%	82,64%	83,70%
2	78,69%	87,43%	58,20%	68,75%	66,91%	84,55%	82,26%	78,88%	81,70%
3	97,85%	98,56%	78,69%	98,46%	51,04%	64,06%	61,79%	87,90%	91,18%
4	98,87%	69,34%	90,67%	98,87%	46,60%	50,86%	71,13%	79,71%	80,96%
5	99,42%	95,53%	65,00%	69,40%	57,48%	63,08%	64,23%	86,90%	87,90%
6	98,67%	92,10%	91,79%	100,00%	58,97%	64,25%	66,12%	88,46%	89,81%
7	98,10%	98,93%	77,86%	100,00%	58,57%	75,23%	80,65%	90,67%	94,55%
8	98,76%	95,27%	85,65%	96,95%	66,36%	75,39%	71,91%	90,30%	92,34%
9	97,62%	87,62%	51,67%	59,24%	44,83%	55,91%	73,54%	81,84%	83,70%
10	95,93%	78,48%	80,73%	84,62%	61,83%	65,34%	65,89%	81,89%	82,63%
11	98,72%	92,63%	82,05%	87,07%	72,63%	77,73%	74,82%	89,92%	90,94%
12	98,83%	77,13%	71,49%	83,58%	38,12%	45,83%	68,04%	79,35%	81,33%
13	95,53%	94,73%	85,36%	95,33%	62,88%	70,59%	83,75%	89,79%	91,56%
14	63,48%	87,50%	72,80%	79,50%	66,51%	74,21%	77,10%	74,48%	75,92%

Tabela C.2 – Tabela de dados por vídeo

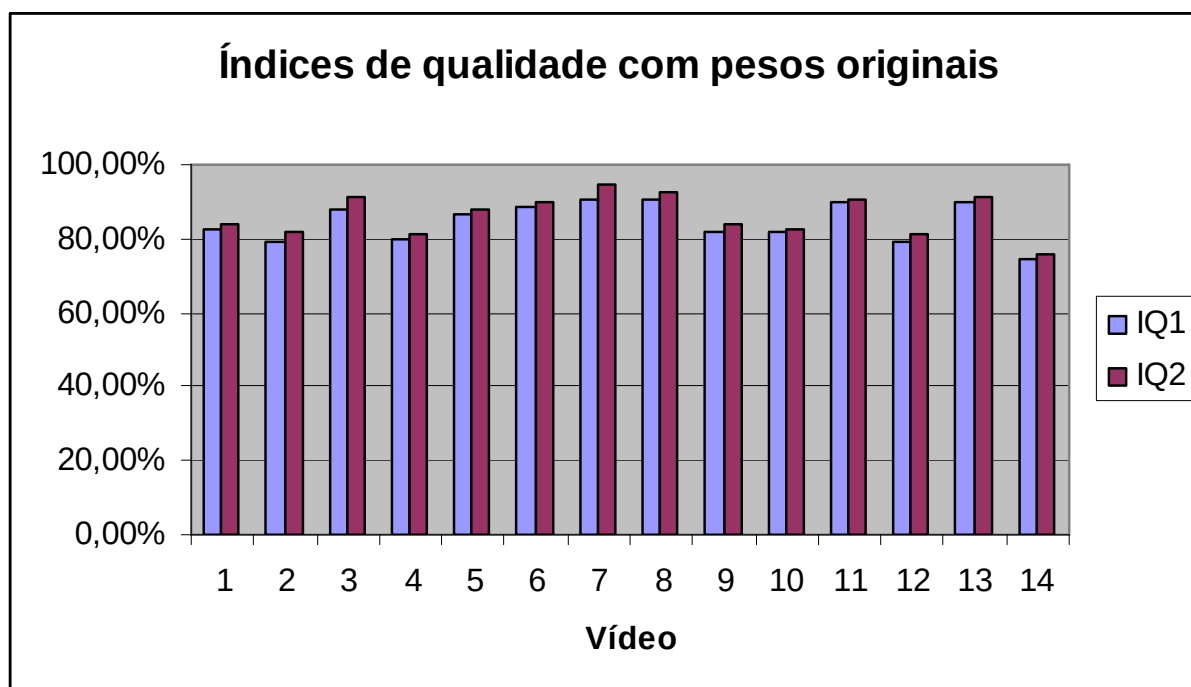


Figura C.1 – Índices de qualidade

Característica	Peso
Campo	20%
Adjacencia	20%
JSC	20%
JCC	20%
RCCP	20%

Tabela C.3 – Tabela de pesos

Vídeo	CMPO	ADJ	JSC1	JSC2	JCC1	JCC2	RCCP	IQ1	IQ2
1	98,32%	80,75%	76,92%	82,47%	58,62%	63,75%	64,06%	75,73%	77,87%
2	78,69%	87,43%	58,20%	68,75%	66,91%	84,55%	82,26%	74,70%	80,34%
3	97,85%	98,56%	78,69%	98,46%	51,04%	64,06%	61,79%	77,59%	84,15%
4	98,87%	69,34%	90,67%	98,87%	46,60%	50,86%	71,13%	75,32%	77,81%
5	99,42%	95,53%	65,00%	69,40%	57,48%	63,08%	64,23%	76,33%	78,33%
6	98,67%	92,10%	91,79%	100,00%	58,97%	64,25%	66,12%	81,53%	84,23%
7	98,10%	98,93%	77,86%	100,00%	58,57%	75,23%	80,65%	82,82%	90,58%
8	98,76%	95,27%	85,65%	96,95%	66,36%	75,39%	71,91%	83,59%	87,66%
9	97,62%	87,62%	51,67%	59,24%	44,83%	55,91%	73,54%	71,05%	74,79%
10	95,93%	78,48%	80,73%	84,62%	61,83%	65,34%	65,89%	76,57%	78,05%
11	98,72%	92,63%	82,05%	87,07%	72,63%	77,73%	74,82%	84,17%	86,20%
12	98,83%	77,13%	71,49%	83,58%	38,12%	45,83%	68,04%	70,72%	74,68%
13	95,53%	94,73%	85,36%	95,33%	62,88%	70,59%	83,75%	84,45%	87,99%
14	63,48%	87,50%	72,80%	79,50%	66,51%	74,21%	77,10%	73,48%	76,36%

Tabela C.4 – Tabela de dados por vídeo

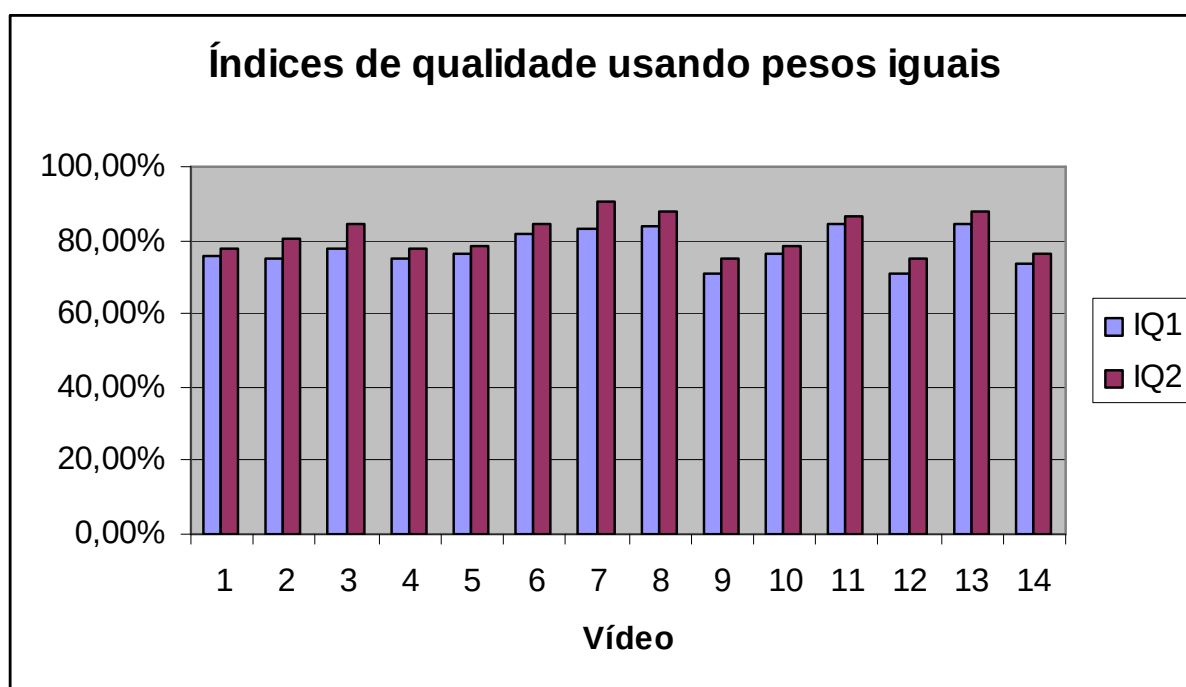


Figura C.2 – Índices de qualidade

Característica	Peso
Campo	10%
Adjacencia	10%
JSC	30%
JCC	30%
RCCP	20%

Tabela C.5 – Tabela de pesos

Vídeo	CMPO	ADJ	JSC1	JSC2	JCC1	JCC2	RCCP	IQ1	IQ2
1	98,32%	80,75%	76,92%	82,47%	58,62%	63,75%	64,06%	71,38%	74,59%
2	78,69%	87,43%	58,20%	68,75%	66,91%	84,55%	82,26%	70,60%	79,05%
3	97,85%	98,56%	78,69%	98,46%	51,04%	64,06%	61,79%	70,92%	80,76%
4	98,87%	69,34%	90,67%	98,87%	46,60%	50,86%	71,13%	72,23%	75,96%
5	99,42%	95,53%	65,00%	69,40%	57,48%	63,08%	64,23%	69,08%	72,08%
6	98,67%	92,10%	91,79%	100,00%	58,97%	64,25%	66,12%	77,53%	81,58%
7	98,10%	98,93%	77,86%	100,00%	58,57%	75,23%	80,65%	76,76%	88,40%
8	98,76%	95,27%	85,65%	96,95%	66,36%	75,39%	71,91%	79,39%	85,49%
9	97,62%	87,62%	51,67%	59,24%	44,83%	55,91%	73,54%	62,18%	67,78%
10	95,93%	78,48%	80,73%	84,62%	61,83%	65,34%	65,89%	73,39%	75,61%
11	98,72%	92,63%	82,05%	87,07%	72,63%	77,73%	74,82%	80,50%	83,54%
12	98,83%	77,13%	71,49%	83,58%	38,12%	45,83%	68,04%	64,09%	70,03%
13	95,53%	94,73%	85,36%	95,33%	62,88%	70,59%	83,75%	80,25%	85,55%
14	63,48%	87,50%	72,80%	79,50%	66,51%	74,21%	77,10%	72,31%	76,63%

Tabela C.6 – Tabela de dados por vídeo

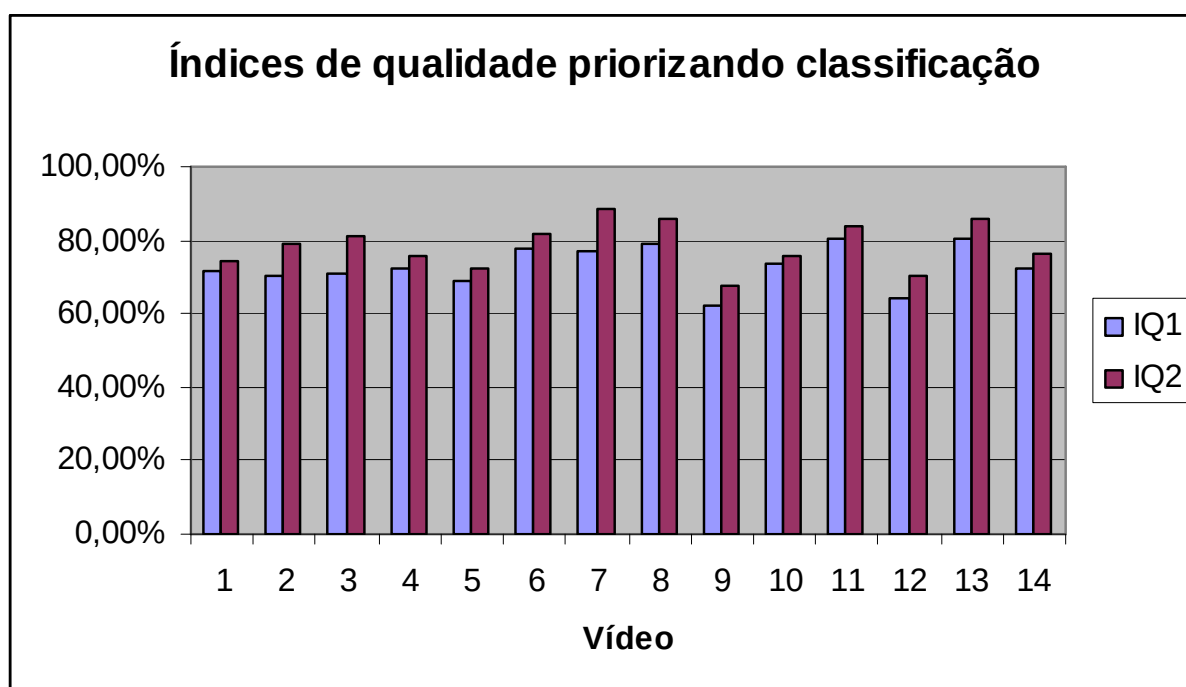


Figura C.3 – Índices de qualidade

Característica	Peso
Campo	0%
Adjacencia	0%
JSC	35%
JCC	35%
RCCP	30%

Tabela C.7 – Tabela de pesos

Vídeo	CMPO	ADJ	JSC1	JSC2	JCC1	JCC2	RCCP	IQ1	IQ2
1	98,32%	80,75%	76,92%	82,47%	58,62%	63,75%	64,06%	66,66%	70,40%
2	78,69%	87,43%	58,20%	68,75%	66,91%	84,55%	82,26%	68,47%	78,33%
3	97,85%	98,56%	78,69%	98,46%	51,04%	64,06%	61,79%	63,94%	75,42%
4	98,87%	69,34%	90,67%	98,87%	46,60%	50,86%	71,13%	69,38%	73,74%
5	99,42%	95,53%	65,00%	69,40%	57,48%	63,08%	64,23%	62,14%	65,63%
6	98,67%	92,10%	91,79%	100,00%	58,97%	64,25%	66,12%	72,61%	77,32%
7	98,10%	98,93%	77,86%	100,00%	58,57%	75,23%	80,65%	71,94%	85,52%
8	98,76%	95,27%	85,65%	96,95%	66,36%	75,39%	71,91%	74,78%	81,89%
9	97,62%	87,62%	51,67%	59,24%	44,83%	55,91%	73,54%	55,84%	62,37%
10	95,93%	78,48%	80,73%	84,62%	61,83%	65,34%	65,89%	69,66%	72,25%
11	98,72%	92,63%	82,05%	87,07%	72,63%	77,73%	74,82%	76,58%	80,13%
12	98,83%	77,13%	71,49%	83,58%	38,12%	45,83%	68,04%	58,78%	65,71%
13	95,53%	94,73%	85,36%	95,33%	62,88%	70,59%	83,75%	77,01%	83,20%
14	63,48%	87,50%	72,80%	79,50%	66,51%	74,21%	77,10%	71,89%	76,93%

Tabela C.8 – Tabela de dados por vídeo

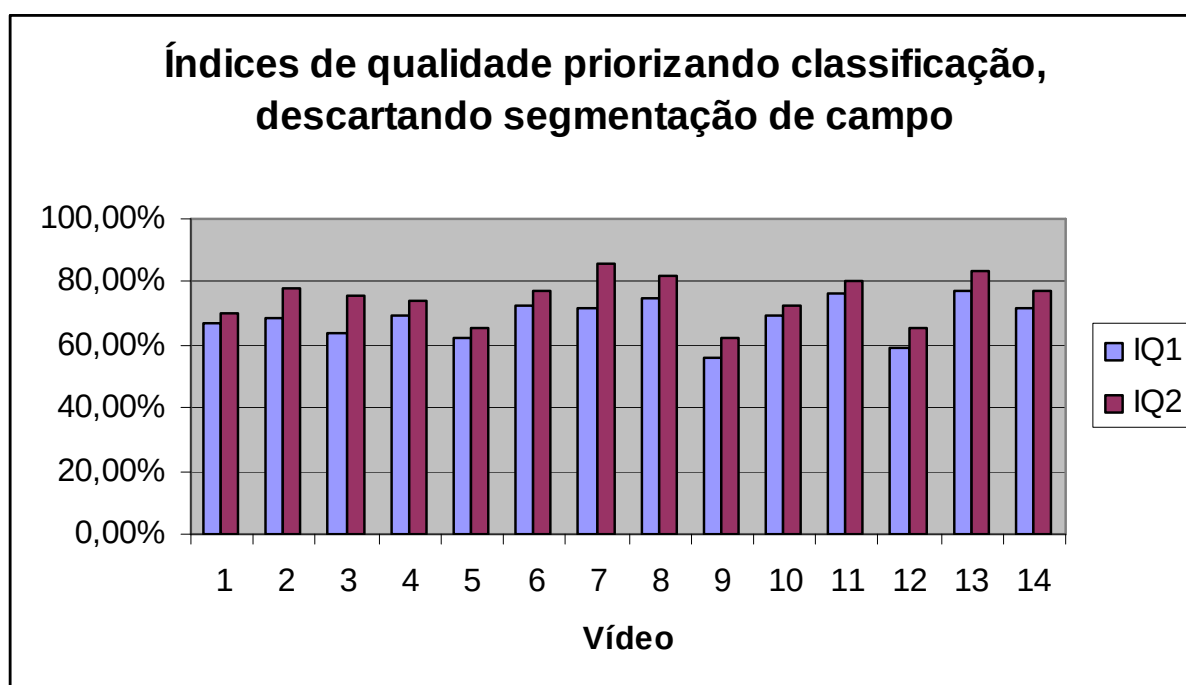


Figura C.4 – Índices de qualidade

Característica	Peso
Campo	50%
Adjacencia	50%
JSC	0%
JCC	0%
RCCP	0%

Tabela C.9 – Tabela de pesos

Vídeo	CMPO	ADJ	JSC1	JSC2	JCC1	JCC2	RCCP	IQ1	IQ2
1	98,32%	80,75%	76,92%	82,47%	58,62%	63,75%	64,06%	89,54%	89,54%
2	78,69%	87,43%	58,20%	68,75%	66,91%	84,55%	82,26%	83,06%	83,06%
3	97,85%	98,56%	78,69%	98,46%	51,04%	64,06%	61,79%	98,21%	98,21%
4	98,87%	69,34%	90,67%	98,87%	46,60%	50,86%	71,13%	84,10%	84,10%
5	99,42%	95,53%	65,00%	69,40%	57,48%	63,08%	64,23%	97,48%	97,48%
6	98,67%	92,10%	91,79%	100,00%	58,97%	64,25%	66,12%	95,39%	95,39%
7	98,10%	98,93%	77,86%	100,00%	58,57%	75,23%	80,65%	98,52%	98,52%
8	98,76%	95,27%	85,65%	96,95%	66,36%	75,39%	71,91%	97,01%	97,01%
9	97,62%	87,62%	51,67%	59,24%	44,83%	55,91%	73,54%	92,62%	92,62%
10	95,93%	78,48%	80,73%	84,62%	61,83%	65,34%	65,89%	87,21%	87,21%
11	98,72%	92,63%	82,05%	87,07%	72,63%	77,73%	74,82%	95,68%	95,68%
12	98,83%	77,13%	71,49%	83,58%	38,12%	45,83%	68,04%	87,98%	87,98%
13	95,53%	94,73%	85,36%	95,33%	62,88%	70,59%	83,75%	95,13%	95,13%
14	63,48%	87,50%	72,80%	79,50%	66,51%	74,21%	77,10%	75,49%	75,49%

Tabela C.10 – Tabela de dados por vídeo

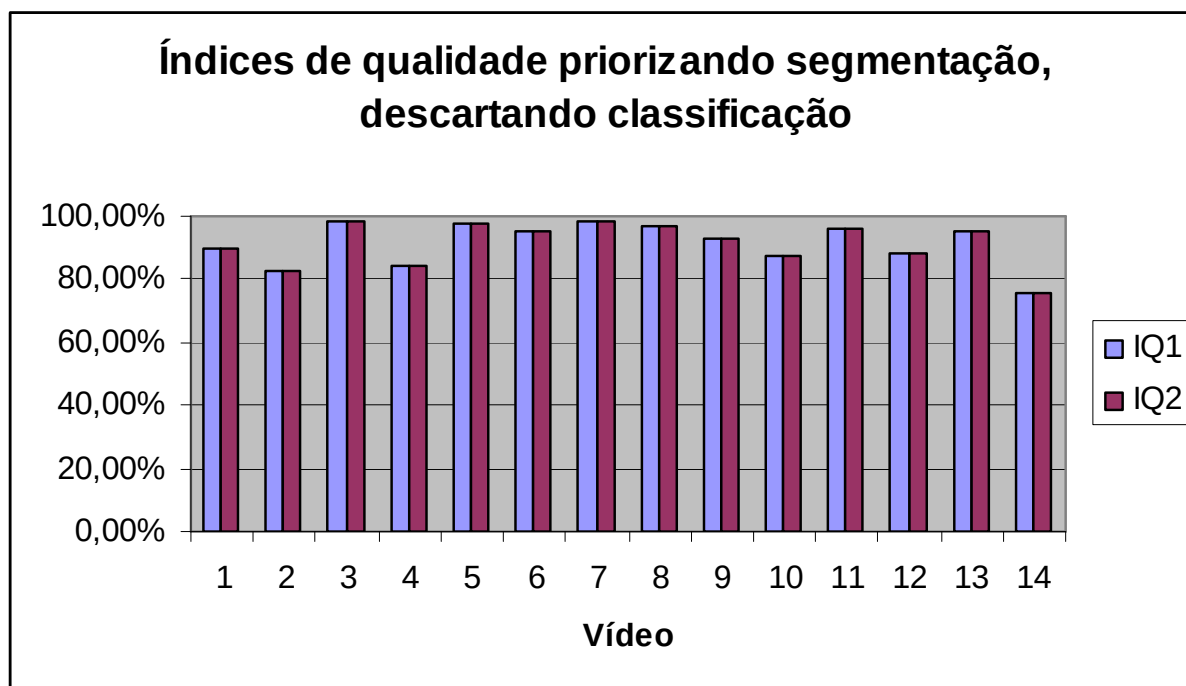


Figura C.5 – Índices de qualidade