

UNIVERSIDADE FEDERAL DE CAMPINA GRANDE

CENTRO DE ENGENHARIA ELÉTRICA E INFORMÁTICA

Coordenação de Pós-Graduação em Ciência da Computação

DISSERTAÇÃO DE MESTRADO

**AVALIAÇÃO DE EXTRAÇÃO DE CARACTERÍSTICAS ESPACIAIS PARA ANÁLISE
DE VÍDEOS DE FUTEBOL**

PAULO DE TARSO FIRMINO JÚNIOR

ORIENTADORES:

HERMAN MARTINS GOMES

JOÃO MARQUES DE CARVALHO

CAMPINA GRANDE, PARAÍBA

FEVEREIRO - 2012

UNIVERSIDADE FEDERAL DE CAMPINA GRANDE

CENTRO DE ENGENHARIA ELÉTRICA E INFORMÁTICA

Coordenação de Pós-Graduação em Ciência da Computação

DETECÇÃO DE ELEMENTOS VISUAIS RELEVANTES EM VÍDEOS DE FUTEBOL POR SEGMENTAÇÃO ESPACIAL

PAULO DE TARSO FIRMINO JÚNIOR

Dissertação submetida à Coordenação do Curso de Pós-Graduação em Ciência da Computação do Centro de Engenharia Elétrica e Informática da Universidade Federal de Campina Grande – Campus I, como parte dos requisitos necessários à obtenção do grau de Mestre em Ciência da Computação (MSc)

Área de concentração: Ciência da Computação

Linha de pesquisa: Metodologia de Técnicas de Computação

Herman Martins Gomes

João Marques de Carvalho

Orientadores

Campina Grande - Paraíba

Fevereiro – 2012



F525a Firmino Junior, Paulo de Tarso
Avaliação de extração de características espaciais para
análise de vídeos de futebol / Paulo de Tarso Firmino
Junior. - Campina Grande, 2012.
89 f. : il.

Dissertação (Mestrado em Ciência da Computação) -
Universidade Federal de Campina Grande, Centro de
Engenharia Elétrica e Informática.

1. Visão Computacional. 2. Sumarização de Vídeos de
Futebol. 3. Processamento de Imagens. I. Gomes, Herman
Martins. II. Carvalho, João Marques de. III. Título

CDU 004.932(043)

**"AVALIAÇÃO DE EXTRAÇÃO DE CARACTERÍSTICAS ESPACIAIS PARA A ANÁLISE
DE VÍDEOS DE FUTEBOL"**

PAULO DE TARSO FIRMINO JÚNIOR

DISSERTAÇÃO APROVADA EM 28/02/2012



JOÃO MARQUES DE CARVALHO, Ph.D
Orientador(a)



HERMAN MARTINS GOMES, Ph.D
Orientador(a)



JOSÉ EUSTAQUIO RANGEL DE QUEIROZ, D.Sc
Examinador(a)

DÍBIO LEANDRO BORGES, Ph.D
Examinador(a)

CAMPINA GRANDE - PB



UNIVERSIDADE DE BRASÍLIA
INSTITUTO DE CIÊNCIAS EXATAS
DEPARTAMENTO DE CIÊNCIA DA COMPUTAÇÃO
70910-900 - Brasília - DF - Brasil
Caixa Postal: 4466
Tel.: 55 61 3107 6407

Brasília, 28 de Fevereiro de 2012.

À: Coordenação do Programa de Pós-graduação em Ciência da Computação da UFCG

Prezados(as) Srs.(as),

Declaro, por intermédio desta, que participei da Banca Examinadora do aluno Paulo de Tarso Firmino Júnior, no dia 28 de Fevereiro de 2012, de 08h30 às 11h00, por vídeoconferência (através do skype). O trabalho foi considerado "Aprovado", e comentários e pequenas correções passadas ao orientador Prof.Dr.Herman Martins Gomes. Agradeço o convite pela participação, deixando meus votos de apreço e consideração.

Atenciosamente,


Prof. Dr.Dívio Leandro Borges

AGRADECIMENTOS

Primeiramente, a Deus, pela saúde e oportunidade de realizar este trabalho.

A toda minha família, pelo apoio. Em especial ao meu pai, Paulo de Tarso, principal incentivador e motivador desta conquista, minha mãe, Maria Elizabete, pelas palavras de carinho e apoio em todos os momentos, meus irmãos, Pedro Paulo e Priscila, que mesmo perto ou longe tiveram sua importância. E aos meus avôs que se foram nesse período: Vô Antônio V-8 e Vovô Pedro "arrumadinho" que certamente, onde estiverem, estarão felizes com mais esta conquista.

À minha "galega" Hanna Raphaelli, minha noiva e futura esposa, que tanto me deu forças para concluir este trabalho, me apoiou, ficou ao meu lado, sendo minha válvula de escape nos momentos difíceis e entendeu sempre que não era possível sair e/ou se ver por conta das atribuições do mestrado.

Aos orientadores Herman Martins Gomes e João Marques de Carvalho, pela dedicação e paciência durante a orientação.

Aos companheiros do Laboratório de Visão Computacional: Eanes, Cláudio, Xhycko, Eduardo, Catharine, Allyson e Daniel.

Aos mais que companheiros de mestrado e amigos Vinícius Porto e Eduardo Moura que, desde o começo, estiveram ao meu lado nos trabalhos das disciplinas e virando noite para as provas.

À Nízia pela consultoria, sua disponibilidade e atenção nas correções foram fundamentais para o sucesso deste trabalho.

A todos que fazem a COPIN, dentre professores e funcionários, em especial Aninha, Vera e Rebeka, sempre prontas a atender aos pedidos urgentíssimos, em prazos curtíssimos.

À CAPES pelo apoio financeiro.

RESUMO

A popularização do futebol, bem como o advento da TV digital, têm despertado o interesse por técnicas automáticas de sumarização de vídeos de futebol. Neste contexto, o presente trabalho objetiva investigar técnicas existentes e propor novas técnicas destinadas aos problemas de Segmentação do Gramado, Segmentação das Linhas do Campo e Classificação de Tomadas. Uma base de dados de quadros de várias partidas de futebol foi especialmente construída para a avaliação experimental das técnicas investigadas. Trabalhos relacionados na área de Segmentação do Gramado, em geral, fundamentam-se na noção de cor dominante e em heurísticas, mas carecem de uma validação experimental objetiva. Neste trabalho, foi proposta uma estratégia de *template matching*, combinada à técnica de vizinhos mais próximos, para a segmentação do gramado. Uma validação dos resultados, com o auxílio de curvas ROC, demonstrou que os resultados obtidos a partir do método proposto foram superiores àqueles obtidos a partir de um conjunto de abordagens existentes, atingindo 98,32% de Taxa de Positivos Verdadeiros (TPR) e 16,08% de Taxa de Positivos Falsos (FPR), na melhor configuração de parâmetros. Os resultados dos trabalhos revisados sobre Segmentação das Linhas do Campo fazem uso da Transformada de *Hough*, porém omitem detalhes relevantes ou utilizam algoritmos muito simples. Uma descrição dos parâmetros utilizados na transformada de *Hough*, bem como o desenvolvimento de melhorias no tocante às etapas de pré e pós-processamento, constituem contribuições desta dissertação no aspecto da segmentação das linhas. Na melhor configuração, os resultados de TPR e FPR foram 77,30% e 6,23%, respectivamente, os quais são passíveis de melhorias em futuras extensões, a partir do emprego da transformada de *Hough* Circular. Para a Classificação de Tomadas, fez-se uso de uma rede neural MLP (*Multi Layer Perceptron*), recebendo diversos atributos derivados das etapas anteriores, ao passo que trabalhos relacionados não consideraram características relevantes, tais como a quantidade de linhas e o número de *blobs*. A validação dos resultados neste problema também foi obtida por meio de curvas ROC, atingindo como melhores resultados TPR de 99,60% e FPR de 0,16%.

ABSTRACT

The popularization of soccer as well as the advent of the digital TV has awakened the interest for automatic soccer video summarization techniques. Within this context, this work aims to investigate existing techniques and to propose new ones applied to the problems of Grass Segmentation, Field Lines Segmentation and Video Shot Classification. A database of video sequences from several matches has been specially built for the experimental evaluation of the investigated techniques. Related work in the area of Grass Segmentation, in general, is based on the notion of dominant color and on some heuristics, but it lacks an objective statistical evaluation. In this work, a strategy is proposed based on template matching combined with the nearest neighborhoods technique in order to perform grass segmentation. A validation of the results, performed with the help of ROC curves, has shown that the results of the proposed method were superior to those of a set of existing approaches, reaching 98.32% of True Positive Rate (TPR) and 16.08% of False Positive Rate (FPR), within the best configuration of parameters. The reviewed work in Field Line Segmentation employs the Hough Transform, but omits relevant details or applies very simple algorithms. A description of the parameters used in Hough transform as well as the development of improvements to pre and post-processing steps were contributions of this dissertation with regard to the line segmentation problem. In the best configuration, the TPR and FPR results were 77.30% and 6.23%, respectively, which may be improved in future extensions by means of the Circular Hough Transform. For the Video Shot Classification, a Multilayer Perceptron neural network has been employed, which received as input features derived from the previous steps, whereas related work did not consider relevant features, as the amount of detected lines and the number of blobs. Result validation in this case was also obtained by means of ROC curves, yielding as best results TPR and FPR of 99.60% and 0.16%.

CONTEÚDO

CAPÍTULO 1	1
INTRODUÇÃO	1
1.1. MOTIVAÇÃO	3
1.2. OBJETIVOS E RELEVÂNCIA	4
1.3. ESTRUTURA DA DISSERTAÇÃO	5
CAPÍTULO 2	8
TRABALHOS RELACIONADOS	8
2.1. SEGMENTAÇÃO DO GRAMADO E CLASSIFICAÇÃO DE TOMADAS	9
2.2. DETECÇÃO DE BOLA	14
2.3. DETECÇÃO DE JOGADORES	18
2.4. SEGMENTAÇÃO DAS LINHAS DO CAMPO	23
2.5. CONSIDERAÇÕES FINAIS DO CAPÍTULO	29
CAPÍTULO 3	31
BASE DE IMAGENS	31
3.1. SEGMENTAÇÃO DA GRAMA E DAS LINHAS	31
3.2. CLASSIFICAÇÃO DE TOMADAS	36
3.3. CONSIDERAÇÕES FINAIS DO CAPÍTULO	36
CAPÍTULO 4	38
ABORDAGENS PROPOSTAS	38
4.1. SEGMENTAÇÃO DO GRAMADO	38
4.2. SEGMENTAÇÃO DAS LINHAS DO CAMPO	41
4.3. CLASSIFICAÇÃO DE TOMADAS	48
4.4. DETALHES DE IMPLEMENTAÇÃO	50
4.4.1. SEGMENTAÇÃO DAS LINHAS DO CAMPO	50
4.4.2. CLASSIFICAÇÃO DE TOMADAS	52
4.5. CONSIDERAÇÕES FINAIS DO CAPÍTULO	54
CAPÍTULO 5	57
AVALIAÇÃO EXPERIMENTAL	57
5.1. SEGMENTAÇÃO DO GRAMADO	57

5.2. SEGMENTAÇÃO DE LINHAS DO CAMPO	63
5.3. CLASSIFICAÇÃO DE TOMADAS	66
5.4. CONSIDERAÇÕES FINAIS DO CAPÍTULO	67
CAPÍTULO 6	69
CONSIDERAÇÕES FINAIS.....	69
6.1. CONTRIBUIÇÕES DO TRABALHO REALIZADO	69
6.2. TRABALHOS FUTUROS	70
REFERÊNCIAS BIBLIOGRÁFICAS.....	72
APÊNDICE A	77
APÊNDICE B	84
APÊNDICE C	87

LISTA DE TABELAS

Tabela 2.1 – Sumário de técnicas revisadas para a Detecção de Bola.....	18
Tabela 2.2 – Sumário de técnicas revisadas para a Detecção de Jogadores.	22
Tabela 2.3 – Sumário de técnicas revisadas para a Segmentação das Linhas do Campo	29
Tabela 3.1 – Relação entre elementos e cores.....	35
Tabela 3.2 – Características da Base de Imagens (Segmentação das Linhas)	357
Tabela 3.3 – Características da Base de Imagens criada para Classificação de tomadas.	38
Tabela 5.1 – Métricas utilizadas na análise ROC.....	58
Tabela 5.2 – Melhores Resultados da Segmentação de Linhas.	65
Tabela 5.3 – Resultados da Classificação das Tomadas	67
Tabela A.1 – Imagens originais da base e imagens segmentadas pela abordagem proposta.....	77
Tabela B.1 – Segmentação das Linhas do Campo por meio da abordagem proposta.	84
Tabela C.1 – Imagens “Positivos Falsos” na Classificação de tomadas.....	87
Tabela C.2 – Imagens “Negativos Falsos” na Classificação de tomadas	88

LISTA DE FIGURAS

Figura 1.1 – Arquitetura para desenvolvimento de um sistema de análise de vídeos de futebol.	6
Figura 2.1 – Exemplos de histogramas para diferentes tipos de tomadas: (a) e (b) Imagens de longa distância do campo de futebol; (c) Imagem de média distância. Abaixo de cada imagem, o respectivo histograma em nível de cinza.	11
Figura 2.2 – Transformada Circular de <i>Hough</i>: (a) Imagem Original; (b) Acumulador de Saída da Transformada de <i>Hough</i>; (c) Detecção do Círculo na Imagem.	15
Figura 2.3 – Ilustração do método de Figueroa, Leite e Barros (2006): (a) Quadro de uma partida de futebol; (b) Background da imagem; (c) Blobs na imagem segmentada.	19
Figura 2.4 – Esquema para detecção de jogadores utilizando por Arnaut e Costa (2009).	21
Figura 2.5 – Estrutura da abordagem proposta em Lotfi e Pourreza (2006).	25
Figura 2.6 – Abordagem de Lofti e Pourreza (2006): (a) Tomada do tipo 1; (b) Quadro após processamento; (c) Tomada do tipo 2; (d) Quadro após processamento.	26
Figura 2.7 – Resultado do método de Sun e Liu (2009): (a) Quadro de uma partida de futebol; (b) Background da imagem; (c) Blobs na imagem segmentada; (d) Imagem após a aplicação do Thinning.	28
Figura 4.1 – Resultado da etapa de Segmentação do Gramado: (a) Imagem original; (b) Imagem que estima a probabilidade do pixel não ser Grama; (c) Imagem segmentada após aplicação do limiar.	40
Figura 4.2 – Tentativa de segmentação das Linhas utilizando abordagem similar ao método de Firmino Jr. e Gomes (2007): (a) Imagem Original; (b) Imagem resultante.	42
Figura 4.3 – Aplicação das máscaras para esqueletização: (a) Imagem Original; (b) Imagem resultante; (c) Histogramas da Imagem resultante.	44
Figura 4.4 – Processo de Binarização: (a) Imagem original; (b) Imagem resultante.	45
Figura 4.5 – Segmentação de Linhas com falhas: (a) Imagem original; (b) Imagem resultante.	46

Figura 4.6 – Eliminação das linhas redundantes: (a) Imagem Original; (b) Imagem resultante.....	46
Figura 4.7 – Algoritmo de Eliminação de Linhas Redundantes.....	47
Figura 4.8 – Interface desenvolvida para auxiliar a escolha dos parâmetros da Transformada de <i>Hough</i>.	51
Figura 4.9 – Detecção de Linhas após o algoritmo de eliminação das linhas redundantes.	52
Figura 4.10 – Ferramenta <i>Weka</i> realizando a seleção de atributos mais importantes.....	54
Figura 4.11 – Ferramenta <i>Weka</i> realizando o treinamento da rede neural.	55
Figura 5.1 – Curva ROC para a variação dos k-vizinhos mais próximos.....	59
Figura 5.2 – Curva ROC dos quatro algoritmos abordados.....	60
Figura 5.3 – Segmentação por intermédio do algoritmo Firmino Jr. E Gomes (2007) modificado: (a) Imagem original; (b) Imagem resultante.	61
Figura 5.4 – Segmentação por intermédio do algoritmo Cardoso (2009): (a) Imagem original; (b) Imagem resultante.....	61
Figura 5.5 – Segmentação por intermédio do algoritmo Ekin, Tekalp E Mehrotra (2003): (a) Imagem original; (b) Imagem resultante.	61
Figura 5.6 – Segmentação por intermédio do algoritmo Pallavi et al. (2008): (a) Imagem original; (b) Imagem resultante.....	62
Figura 5.7 – Segmentação das Linhas: (a) Imagem original; (b) Imagem rotulada; (c) Imagem segmentada.	63
Figura 5.8 – Mapa de Erros da Figura 5.7.	64
Figura 5.9 – Curva ROC para a Segmentação das Linhas.	65

LISTA DE SIGLAS E ABREVIATURAS

GT	Ground-Truth
HSV	Hue, Saturation, Value
HSI	Hue, Saturation, Intensity
JPEG	Joint Photographic Experts Group
OpenCV	Open Source Computer Vision (Intel Library)
PCA	Principal Component Analysis
PDI	Processamento Digital de Imagens
RGB	Red, Green, Blue
ROC	Receiver Operating Characteristics
ROI	Region Of Interest
NTSC	National Television System(s) Committee
YIQ	Y (luminância) I e Q (componentes cromáticos em: fase e quadratura)
HD	Hard Disk
FIFA	Fédération Internationale de Football Association
MLP	Multi Layer Perceptron
CHT	Circular Hough Transform
UEFA	Union of European Football Associations
TP	True Positive
FP	False Positive
FN	False Negative
TN	True Negative
EKF	Extended Kalman Filter
LB	Limiar de Binarização
IDE	Integrated Development Environment
TPR	True Positive Rate
FPR	False Positive Rate

Capítulo 1

Introdução

A popularização de vídeos na *Internet* se torna cada dia maior, devido, sobretudo, ao surgimento de ferramentas *on-line* que permitem a exibição e o compartilhamento de vídeos, tais como YOUTUBE, GOOGLE VIDEOS, MEGA VIDEO e VIMEO. O YOUTUBE, fundado em 2005 pelos três pioneiros do *PayPal* (empresa que acata a transferência de dinheiro entre indivíduos ou negociantes), utiliza o formato *Adobe Flash* para a exibição diária de cerca de 100 milhões de vídeos, além da postagem de outros 65 mil (GIESSEL, 2009). Por sua vez, o GOOGLE VIDEOS, que está chegando ao seu fim (PEPERAIO, 2011), é um serviço web que faz a busca dos mais diversos vídeos na *Internet* e disponibiliza *links* para *sites*, como o próprio YOUTUBE. O MEGA VIDEO e o VIMEO utilizam uma abordagem de repositório de vídeos, assim como o YOUTUBE.

Outro fator importante para a popularização de vídeos foi o surgimento da TV Digital em rede aberta no Brasil, em 2 de Dezembro de 2007 (SANTOS, 2007), assim como a popularização das empresas de TV por assinatura, das quais só a Sky Brasil possui cerca de 9,3 milhões de assinantes (SILVA, 2009). Paralelamente a este crescimento, há o surgimento de diversos aparelhos responsáveis pela reprodução e gravação de vídeos digitais nos mais diversos formatos (*avi, mpeg, rmvb*), e.g., Asus O!Play HDP-R1 HD Media Player (CHESTER, 2010). A partir destes fatos, não é difícil entender o interesse de empresas e pesquisadores em investigar e desenvolver técnicas de sumarização automática de vídeos, ou seja, a extração de partes importantes do vídeo, sem que haja perda de seu conteúdo semântico (YE ET AL., 2005).

As técnicas de sumarização de vídeos podem ser classificadas em duas categorias principais: *keyframe* e *video skim*. As técnicas *keyframe* utilizam a extração de um conjunto de quadros-chave do vídeo original, que resulta em resumos estáticos. Por outro lado, as técnicas da classe *video skim* produzem um resumo dinâmico, que agrupa um conjunto de tomadas a partir da análise de similaridade ou da relação temporal entre os quadros (TRUONG; VENKATESH, 2007).

Com a popularização dos vídeos digitais e o aumento da quantidade de vídeos produzidos diariamente, é evidente que diversos segmentos poderiam ser beneficiados pela sumarização de aspectos de interesse, tais como empresas de televisão que possuem uma grande quantidade de mídia digital em seus acervos, sistemas de monitoramento de segurança, públicos ou privados, que produzem diariamente uma quantidade substancial de vídeos ou até mesmo os usuários de câmeras digitais, cada vez mais acessíveis, os quais produzem seus próprios vídeos (ASSFALG ET AL., 2003).

Pesquisas realizadas pela FIFA, entidade máxima do futebol mundial, mostraram que aproximadamente 270 milhões de pessoas no mundo estão ativamente envolvidas com o futebol (aproximadamente 4% da população mundial) e que quase metade do planeta estava acompanhando a Copa do Mundo de 2010, realizada na África do Sul (aproximadamente 3,2 bilhões de telespectadores) (PRESS, 2011), o que ratifica a grande popularidade do futebol no cenário mundial e desperta o interesse na sumarização de seus vídeos. Existem diversos elementos significativos em uma partida de futebol, tais como a bola, os jogadores e as linhas do campo que, uma vez sendo detectados e/ou rastreados, podem ser de grande valia para a realização de uma sumarização de qualidade.

A presente dissertação trata do problema de segmentação espacial de partes significativas em uma partida de futebol. Os aspectos espaciais podem ser definidos como aqueles presentes em quadros de forma isolada tais como a grama e as linhas do campo, sem a restrição temporal e levando-se em consideração que o vídeo foi filmado com câmera parada.

Em contrapartida, os aspectos temporais incluem a bola e os jogadores, que necessitam uma sequência de quadros no decorrer do tempo, para o sucesso da detecção. Neste presente trabalho, apenas os aspectos espaciais foram considerados, enquanto as abordagens temporais ficam como trabalhos futuros. Pretende-se realizar uma análise comparativa entre as principais técnicas existentes, visando desenvolver um sistema que agrupe as melhores abordagens de segmentação/classificação, considerando os seguintes aspectos: grama, linhas do campo e tomadas da câmera (longa distância e outras). Desta forma, foram desenvolvidas, durante a dissertação; (i) uma abordagem original com relação à Segmentação do Gramado e Classificação de Tomadas; e (ii) uma abordagem melhorada para a Segmentação das Linhas do Campo. Para ambas as abordagens foi realizada uma avaliação experimental objetiva com a finalidade de validar os métodos desenvolvidos.

1.1. Motivação

Diante do exposto na seção anterior, a principal motivação para o trabalho desenvolvido foi a investigação e implementação de ferramentas automáticas que pudessem vir a servir de base para um sistema de sumarização automática de partidas de futebol. Tal sistema trabalharia com os aspectos espaciais descritos anteriormente, para produzir um resumo que preservasse todos os eventos relevantes de uma partida. Na presente dissertação, apenas a extração ou segmentação dos elementos espaciais foi abordada, ficando os aspectos temporais reservados para trabalhos futuros.

A segmentação/classificação dos elementos apresentados (grama, tomadas e linhas do gramado) é de fundamental importância para uma etapa posterior de sumarização do vídeo de futebol. Vídeos sumarizados podem ser gerados automaticamente, com vistas apenas ao registro de lances específicos, que sejam de interesse do usuário de uma ferramenta que implemente as técnicas pesquisadas nesta dissertação, bem como de lances de determinado jogador, lances de gol e/ou filmagens da torcida, dentre outros.

Como a segmentação, tal como a classificação de grama e de tomadas, não é objetivo final da maioria das aplicações de Visão Computacional e por ser, em geral, executada de forma manual, sua automação, já é, por si, uma grande motivação para o desenvolvimento da pesquisa ora relatada.

No tocante à segmentação do gramado, outra motivação é a inexistência de uma avaliação experimental objetiva nos trabalhos já publicados sobre este assunto. Apenas em Cardoso (2009), algumas métricas são utilizadas como forma de avaliação da qualidade da técnica. Além disto, não foi encontrada, na literatura revisada, uma base de imagens rotuladas que possibilitasse uma avaliação envolvendo as principais técnicas existentes.

A segmentação das linhas surge como auxílio à verificação de lances que ocorram nas proximidades ou no interior da grande área, local em que ocorrem os principais fatos de uma partida. Sendo assim, esta técnica, combinada à classificação de tomadas pode gerar uma ferramenta capaz de reduzir o tamanho do vídeo original (pela eliminação de *replays*, filmagens da torcida, etc.), não comprometendo seu entendimento, visto que os principais lances da partida serão detectados.

1.2. Objetivos e Relevância

Apesar da existência de muitos trabalhos desenvolvidos na área de vídeos esportivos, não foram encontradas análises estatísticas significativas capazes de validar a eficácia dos métodos existentes. Mediante experimentos e testes estatísticos conhecidos, tais como *T de Student*, *Anova* e *Tukey-Kramer* (LEVINE; BERENSON; STEPHAN, 2004), serão verificadas nesta dissertação quais abordagens produzem os melhores resultados.

Portanto, esta dissertação tem como objetivos específicos realizar não apenas um levantamento do estado da arte dos métodos existentes para a análise de vídeos de futebol e implementar técnicas que permitam a segmentação dos aspectos espaciais presentes nos mesmos, mas também

realizar uma avaliação experimental objetiva dos resultados apresentados nos principais trabalhos existentes nas áreas de segmentação do gramado, das linhas do campo e classificação de tomadas, bem como desenvolver algoritmos inovadores.

O objetivo geral desta dissertação, após realizada a verificação experimental objetiva para a comparação dos métodos estudados, é desenvolver um sistema contendo todos os segmentadores/classificadores que apresentam o melhor desempenho. Este sistema utiliza um conjunto de imagens como entrada, podendo ainda ser melhorado para que o próprio vídeo seja utilizado. Uma arquitetura para a metodologia de desenvolvimento utilizada nesta dissertação se encontra representada na Figura 1.1. Primeiramente, os vídeos devem ser particionados em imagens para a realização da Fase 1, enquanto a Fase 2 tem início com a construção de um *Ground-Truth* para a realização de uma análise estatística; na Fase 3 as técnicas mais promissoras poderão ser ainda melhoradas, realimentando a Fase 2; por fim, as melhores técnicas serão agrupadas para a construção do sistema.

Durante a revisão bibliográfica, não foi encontrado qualquer sistema que englobasse a detecção do conjunto de elementos apresentados na Seção 1.2. Um sistema com tais características seria de grande utilidade, tanto para usuários interessados em assistir apenas partes de um vídeo de futebol, como para a indústria televisiva, que poderia utilizar as características extraídas para facilitar a indexação, recuperação e sumarização de seus acervos de vídeos.

1.3. Estrutura da Dissertação

O presente documento é composto por seis capítulos. Neste capítulo, são apresentadas as principais motivações e contribuições do trabalho. No **Capítulo 2**, as principais técnicas para detecção de bola, jogadores, tomadas (média, curta e longa distância), linhas do campo e gramado são apresentadas. Adicionalmente, foi criado um quadro comparativo entre as principais técnicas, com a identificação das mais promissoras.

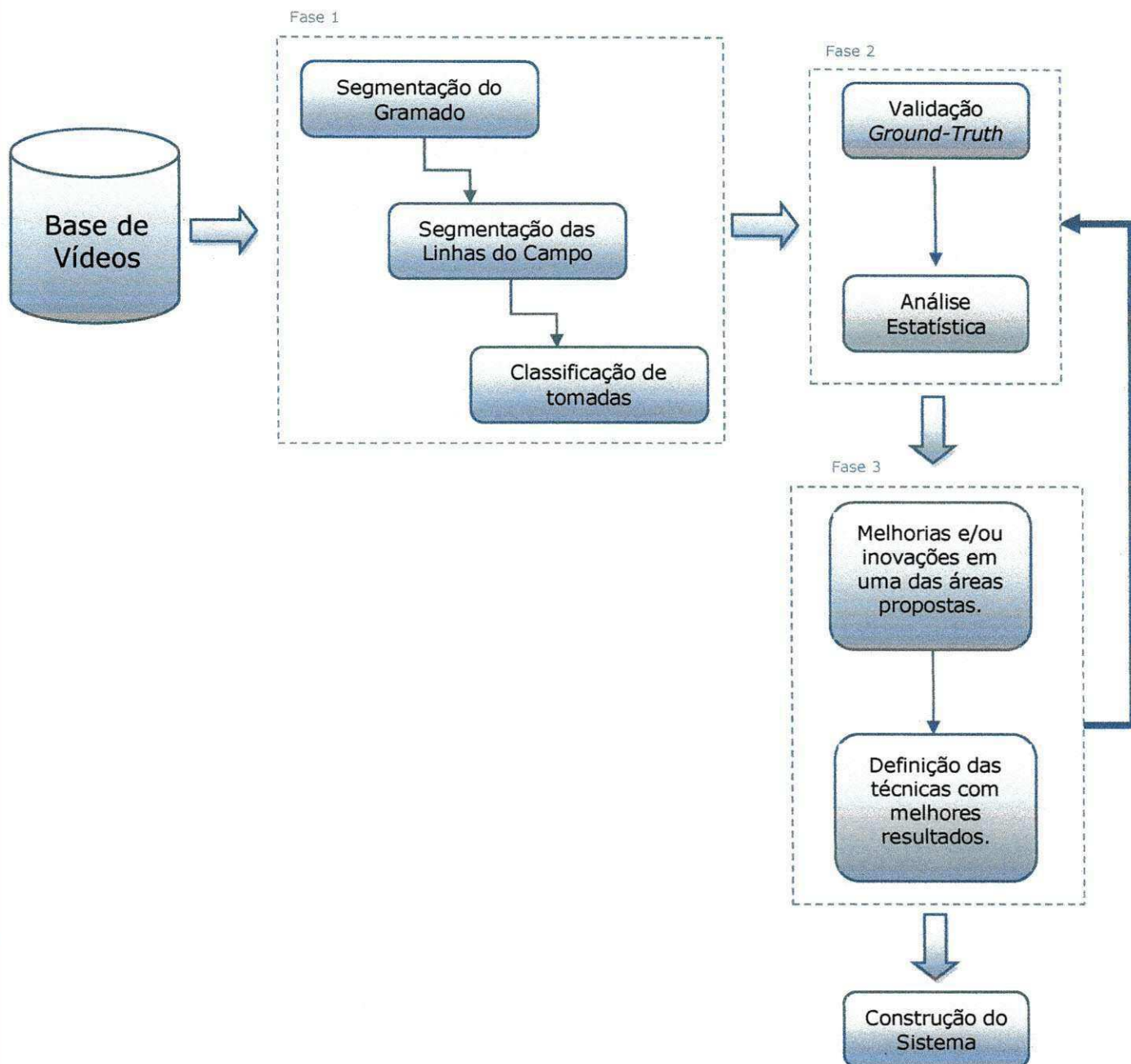


Figura 1.1 – Arquitetura para desenvolvimento de um sistema de análise de vídeos de futebol.

No **Capítulo 3**, se apresenta uma importante contribuição do trabalho desenvolvido, que corresponde à criação de 3 bases de imagens para experimentos. A primeira base é composta por 300 imagens binárias segmentadas em grama e não grama, para serem utilizadas nos testes de Segmentação das Linhas; a segunda base contém 184 imagens segmentadas em jogador, bola, linhas, grama e outros, para testes da Segmentação das Linhas; e, por fim, a terceira base contém 8365 imagens

rotuladas em longa distância (5391) e outras (2974) para testes da Classificação de Tomadas. Estas bases foram construídas a partir de quadros de vídeos da Copa do Mundo 2010, realizada na África do Sul, além de alguns jogos do Campeonato Brasileiro de 2010 e 2011 e da *Champions League* de 2010.

No **Capítulo 4**, são descritas as abordagens propostas para a segmentação do gramado e das linhas do campo, assim como para a classificação de tomadas (longa e outras), sendo apresentado o fluxo de processamento, desde a imagem de entrada até a geração do resultado final. Algumas imagens que representam os passos intermediários são apresentadas e o funcionamento de cada módulo das técnicas propostas é ilustrado.

A avaliação experimental, terceira contribuição central da dissertação, a qual engloba a apresentação de resultados e realização de análises estatísticas, é incluída no **Capítulo 5**. Finalmente, no **Capítulo 6**, são apresentadas as considerações finais resultantes deste estudo, bem como as contribuições e trabalhos futuros que podem ser derivados do tema abordado nesta dissertação.

Capítulo 2

Trabalhos Relacionados

No capítulo anterior, foi apresentada a contextualização do problema abordado na presente dissertação. No presente capítulo, são apresentados e discutidos trabalhos recentes com resultados relevantes na área de segmentação de vídeos de futebol. As abordagens são agrupadas de acordo com a divisão a seguir e apresentadas em ordem cronológica, nas seções subsequentes:

- **Segmentação do Gramado e Classificação de Tomadas:** A classificação de tomadas (curta, média e longa distância) está muito ligada à detecção do gramado, pois uma ocorrência maior ou menor da grama poderá indicar o tipo de tomada, motivo pelo qual os trabalhos estão agrupados em um mesmo tópico.
- **Detecção de Bola:** A bola é o elemento fundamental da partida de futebol. Sua detecção pode trazer informações relevantes, tais como o tempo de posse de bola de cada equipe e o jogador mais participativo na partida, dentre outras.
- **Detecção de Jogadores:** A detecção dos jogadores pode servir, por exemplo, para verificar qual o time que realizou mais ataques e, conjuntamente com a detecção de bola, verificar, de forma automática, o tempo de posse de bola de cada time, dentre outras possíveis aplicações.

- **Segmentação das Linhas do Campo:** As linhas do campo são de fundamental importância, pois definem a área válida do jogo. Sua detecção pode servir, por exemplo, para indicar a saída de bola em determinados lances e também a ocorrência de gol.

2.1. Segmentação do Gramado e Classificação de Tomadas

A transmissão de partidas de futebol é realizada por diversas câmeras com recursos e *zoom*, que permitem alterar o tipo da tomada, de acordo com o interesse no lance. Geralmente, tomadas de longa distância possuem cor dominante (verde) e são utilizadas na apresentação de lances da partida em tempo real; as tomadas de média distância geralmente não possuem cor dominante e são utilizadas para apresentação de *replays* e, por fim, as tomadas de curta distância são filmagens de pessoas específicas da torcida ou de jogadores, associados a algum evento isolado da partida de futebol (SGARBI; BORGES, 2005).

Uma etapa de pré-processamento que muitos autores utilizam para realizar a identificação de tomadas em partidas de futebol é a detecção da cor dominante; neste caso, o verde do gramado. No trabalho de Ekin, Tekalp E Mehrotra (2003), é apresentada uma análise automática em vídeos de partidas de futebol por meio de diversos algoritmos, em que a forma de segmentar a imagem em grama e não grama é realizada por intermédio de uma técnica de detecção de cor dominante.

Em geral, existem grandes variações na tonalidade de verde do gramado, de acordo com o estádio, a iluminação e o horário da partida, dentre outras causas. Como maneira de realizar a detecção de forma consistente, Ekin, Tekalp e Mehrotra (2003) utilizaram histogramas no espaço de cores HSI (*Hue-Saturation-Intensity*). Inicialmente, os autores consideram o pico do histograma em cada canal do espaço de cores HSI e

uma faixa de variação de 20% em torno do pico, para o cálculo do que o autor chamou de cor média, que pode ser considerado o centro de massa do histograma. Após esta etapa, a acromaticidade (valores muito baixos para o canal S) do *pixel* é verificada e, em caso afirmativo, apenas o componente I é utilizado para validar se o *pixel* possui cor dominante; caso contrário, serão considerados os três canais na detecção.

Sgarbi e Borges (2005) propuseram um algoritmo para a segmentação de vídeo, com o propósito de remover as tomadas de média e curta distância, de tal modo que apenas as filmagens de longa distância (que intuitivamente despertam maior interesse) permaneçam. Esta distinção de cenas foi realizada por meio da análise do histograma em níveis de cinza, composto pela média dos canais R, G e B de cada quadro visto. Quando a filmagem é realizada por uma câmera de longa distância, se observam poucas variações no histograma ao longo dos quadros, o qual permanece quase inalterado. Por outro lado, quando ocorre mudança para uma câmera de curta distância, a modificação é facilmente percebida no histograma, conforme pode ser observado na Figura 2.1. Os testes foram realizados com 4 vídeos diferentes dos quais 3 foram realizados no Brasil e 1 no Chile. Não foram apresentados resultados para a Segmentação do Gramado separadamente, apresenta resultados para a detecção de lances significativos, em que a média da taxa de detecção obtida foi 90,8% para um total de 675242 imagens (SGARBI; BORGES, 2005).



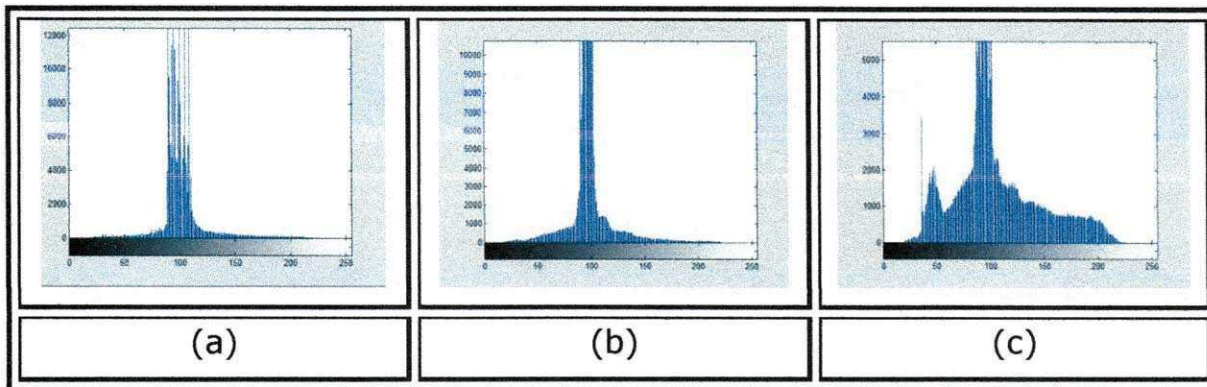


Figura 2.1 – Exemplos de histogramas para diferentes tipos de tomadas: (a) e (b) Imagens de longa distância do campo de futebol; (c) Imagem de média distância. Abaixo de cada imagem, o respectivo histograma em nível de cinza.

Uma vez que a apresentação de tomadas de longa distância em pequenos dispositivos móveis pode reduzir demasiadamente o tamanho da bola, Seo et al. (2007) apresentaram um estudo para suprir este problema, por meio da detecção de regiões de interesse. O trabalho foi dividido em 3 estágios: Aprendizado, Classificação de tomadas e Determinação da ROI (Região de Interesse). O estágio de aprendizado é responsável pela detecção da cor dominante, ou seja, a cor do gramado. Os autores utilizam o modelo cromático HSV e seu histograma acumulado para os 30 primeiros quadros da partida. Primeiramente, Seo et al. (2007) realizaram um teste com o histograma acumulado do canal *Hue*, que deve possuir uma taxa superior ou igual a 70% dos *pixels* no intervalo de valores de cinza 9 a 21, o qual corresponde ao gramado. Caso esta condição não seja verificada, deve-se repetir o experimento com os 30 quadros seguintes. De forma empírica, também foi observado que o pixel será considerado pertencente à classe grama, caso os valores dos canais R, G e B sigam a regra $G > R > B$. Antes da segmentação final da imagem em grama e não grama, deve-se calcular o valor médio para a saturação, a média aritmética simples (*SaturationMean*) e o valor do pico do canal V (*ValuePeakIndex*). A imagem final é segmentada de acordo com a Equação 2.1:

$$\text{Grama}(x, y) = \begin{cases} 1, & \text{se } \left\{ \begin{array}{l} (G > 0,95 \times R) \text{ e } (R > 0,95 \times b) \text{ e} \\ (V(x,y) < 1,25 \times \text{ValuePeakIndex}) \text{ e} \\ (S(x,y) > 0,8 \times \text{SaturationMean}) \end{array} \right. \\ 0, & \text{caso contrário} \end{cases} \quad (2.1)$$

Os resultados apresentados são promissores, uma vez que a média de detecção de tomadas de longa distância e tomadas consideradas outras foram 96,96% e 94,14%, respectivamente, gerando uma média de 96,03% em um total de 5 vídeos.

Outra forma de detecção de grama pode ser vista no trabalho de Pallavi et al. (2008) em que foi utilizado o espaço de cores YIQ, que é utilizado no sistema NTSC (sistema de televisão analógico atualmente utilizado nos Estados Unidos e na maioria dos países da América). Mediante experimentos extensivos com 4500 quadros de 5 diferentes partidas de futebol, da Euro Copa 2004 e UEFA *Champions League* 2003, foi definido um intervalo para os valores de pixel de grama. Com relação ao componente I e caso o *pixel* esteja no intervalo $[I_{min}; I_{max}]$ ele será considerado grama, com $I_{min} = 25$ e $I_{max} = 55$. O mesmo procedimento é seguido para o canal Q, com $Q_{min} = 0$ e $Q_{max} = 12$; caso o valor do *pixel* esteja dentro dos dois intervalos simultaneamente, deverá ser marcado como grama. Após a detecção da cor dominante, é realizada a classificação do quadro, de acordo com o percentual obtido. Caso o quadro possua uma taxa de ocorrência de grama entre 0,75 e 1, será considerado de longa distância; caso esteja entre 0,5 e 0,75, será considerado de média distância e, se a cor dominante do quadro não for o verde (grama), ou seja, não obedeça às condições definidas, o quadro será considerado de curta distância. Os resultados apresentam uma taxa de classificação correta de 96,68% para longa distância, 83,76% para média distância e 87,63% para curta distância, para o total de 5830 quadros de vídeo.

As redes neurais artificiais também surgem como uma técnica promissora na detecção de grama, como se pode verificar no trabalho de Cardoso (2009). Visando automatizar o ajuste de segmentadores baseados em informação de cor e textura, utilizou-se a rede MLP (*Multi Layer Perceptron*) como estrutura de classificação, filtros de Gabor para descrever a textura e um algoritmo de *floodfill* no pós-processamento. Por fim, como forma de melhorar o processamento do algoritmo, foram encontrados os melhores parâmetros dos filtros de Gabor e do algoritmo de *floodfill*, por intermédio de algoritmos genéticos.

O trabalho de Cardoso (2009) apresentou uma diferença importante em relação aos demais em virtude de apresentar uma avaliação experimental objetiva que comprova a obtenção dos melhores resultados. Para a realização dos experimentos foi necessária obtenção de um *Ground-Truth* composto de 25 imagens rotuladas manualmente, disponíveis em Firmino Jr. e Gomes (2007). As medidas utilizadas para realizar a comparação dos métodos existentes foram: Taxa de Detecção e Precisão (*precision*) (resultado da relação entre o número de informações relevantes recuperadas e o total de informações recuperadas, Araújo (1979)) e a Medida-F (média harmônica entre a taxa de detecção e a precisão), que apresentaram taxas de 96,68%, 97,55% e 97,11%, respectivamente.

Durante a revisão bibliográfica, duas dificuldades foram encontradas no tocante à detecção de grama: a existência de poucos trabalhos específicos nesta área e a falta de resultados claros e objetivos nos referidos trabalhos. Grande parte dos trabalhos existentes utiliza a detecção de grama como pré-processamento para a detecção de bola e de jogadores. Por este motivo não é realizada, em geral, uma análise mais significativa dos resultados. Uma exceção é o trabalho de Cardoso (2009). Esta é uma lacuna que a presente dissertação pretende preencher por meio de um estudo comparativo, utilizando a curva ROC como critério de avaliação.

Outra maneira de se realizar a detecção de tomadas em uma partida de futebol é por intermédio da divergência de *Kullback-Leibler* estendida (medida de informação mútua entre quadros consecutivos) como proposta por Lopes, Fukuma e Rodrigues (2011). Este trabalho se concentra no uso de informações temporais do vídeo analisado, em que a divergência de *Kullback-Leibler* faz a comparação de histogramas par-a-par, classificando os quadros em duas classes, a saber: Câmera 1 (longa distância) e Não-Câmera 1 (outras). Com relação aos resultados apresentados, o trabalho de Lopes, Fukuma e Rodrigues (2011) traz um importante diferencial, que é o uso da Curva ROC como forma de obtenção dos parâmetros necessários ao emprego da entropia. Os resultados aparentam ser promissores, tendo em vista que a taxa de detecção está em torno de 97%, para o total de 45000 quadros, em uma hora e meia, de três vídeos diferentes.

O passo seguinte à detecção de grama é a detecção de bola, um dos elementos de maior significado em uma partida de futebol. Os principais trabalhos existentes na área serão apresentados na Seção 2.2, bem como uma análise crítica de cada um deles.

2.2. Detecção de Bola

Conforme comentado anteriormente, um dos objetos dinâmicos mais importantes em uma partida de futebol é a bola. Os lances de maior influência tais como gols, impedimentos e faltas, podem ser verificados por intermédio da detecção e/ou rastreamento da bola. O passo inicial para a detecção de bola, em grande parte dos trabalhos existentes na literatura, é a detecção de grama (apresentada na Seção 2.1), que serve para restringir a área de busca da bola. Depois de detectada a grama, a área da imagem para a localização da bola fica delimitada, podendo melhorar a velocidade de algoritmos e a qualidade de seus resultados.

D’Orazio et al. (2004) buscam a detecção de bola em dois estágios: detecção de candidatos, por meio da transformada circular de *Hough* e inferência dos candidatos, por meio de uma rede neural MLP (*Multi Layer Perceptron*), treinada com exemplos de bola e não bola. A transformada Circular de *Hough* (CHT) é bastante utilizada na literatura como forma de detecção de possíveis candidatos a bola. Ela é uma variação da transformada de *Hough* que utiliza um processo de eleição de círculos, em que os votos são atribuídos aos pontos de passagem dos prováveis círculos existentes na imagem e a detecção de um possível círculo é obtida por intermédio do vetor de acumulação de votos, no qual o valor máximo (cume) representa o centro de um possível círculo (DUARTE, 2003). A Figura 2.2 mostra tal acumulação dos votos.

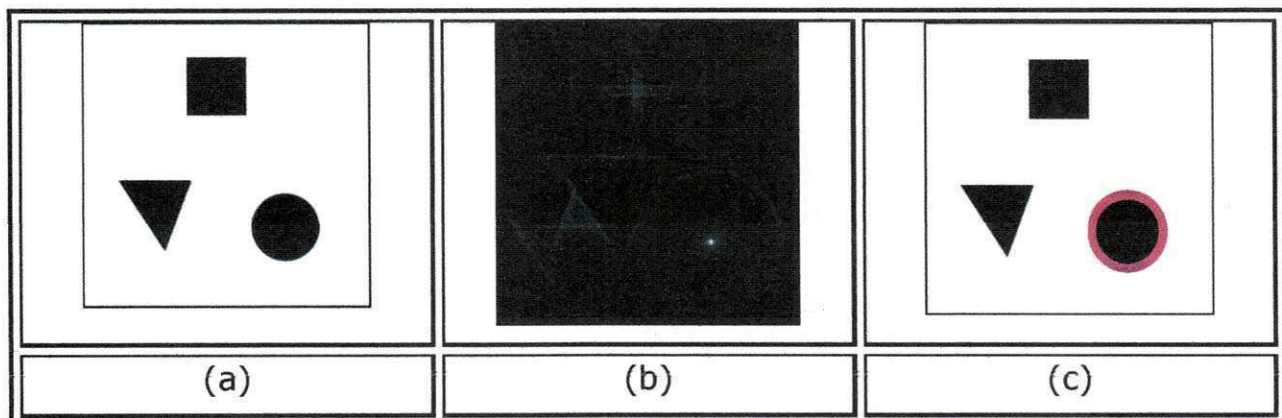


Figura 2.2 – Transformada Circular de *Hough*: (a) Imagem Original; (b) Acumulador de Saída da Transformada de *Hough*; (c) Detecção do Círculo na Imagem.

Ocorrida a detecção de possíveis candidatos a bola é apresentada, em D’Orazio et al. (2004), uma Rede Neural *Multilayer Perceptron* (MLP) com uma camada escondida. Esta arquitetura de Rede Neural foi treinada a partir do algoritmo *backpropagation*, com exemplos de bola e não bola. Um aspecto inovador daquele trabalho é a detecção da bola com a presença de sombra por intermédio do uso de transformadas *Wavelet*. Com relação à avaliação experimental, a taxa de classificação de bola obteve 96,46% de acerto em uma base de teste não mencionada pelo autor.

Outra forma de detectar a bola é apresentada por Yu et al. (2007) que utilizam, além de informações de tamanho e forma, grafos ponderados e o algoritmo de *Viterbi* (SALES, 2004). Primeiro, como em todos os trabalhos

apresentados até este ponto, a detecção da área do jogo é realizada para reduzir o campo de busca do algoritmo; em seguida, é calculada a média das regiões presentes no gramado com a finalidade de reduzir os possíveis candidatos à bola, levando-se em consideração a forma e relacionando a área e o perímetro de cada candidato. Após esta etapa, é realizada a construção de um grafo ponderado, por intermédio dos sucessivos quadros do vídeo, em que os nós correspondem aos candidatos que não foram eliminados, os arcos representam a distância euclidiana entre os respectivos nós nos quadros sucessivos e os pesos dos candidatos são calculados a partir da forma, área e perímetro de cada candidato. Como a bola é o elemento que apresenta maior deslocamento (maior velocidade) entre quadros, fato observado empiricamente por meio da análise de diversos vídeos, o candidato escolhido será aquele que possua a maior soma dos pesos no caminho selecionado. Pelas conclusões apresentadas pelos autores, a utilização do Algoritmo de Viterbi para o rastreamento da bola aparenta resolver o problema proposto, mas apenas resultados experimentais com dois vídeos de duração não especificada, obtidos da Copa de 2006, são indicados no trabalho e nenhuma avaliação estatística foi realizada.

A segmentação de quadros em grama/não grama também pode ser útil na classificação das tomadas, visto que, para cada tipo de tomada, existe uma técnica de detecção de bola mais apropriada. Por exemplo, no trabalho de Pallavi et al. (2008), é utilizada, primeiramente, a técnica CHT (*Circular Hough Transformation*) como forma de identificar possíveis candidatos a bola, tanto em tomadas de longa como em tomadas de média distância. No caso de um quadro de média distância, é utilizada a filtragem dos candidatos segundo o tamanho do raio e posteriormente, é analisada a velocidade da bola ao longo de uma sequência de quadros do vídeo. Caso o quadro seja de longa distância, o algoritmo CHT é utilizado em combinação com um algoritmo de programação dinâmica para a detecção de bola. Esta abordagem é interessante, em virtude de aplicar técnicas distintas, de acordo com o tipo de tomada da imagem e apresentar bons resultados, na detecção de bola, em quadros de longa distância. A taxa de revocação (*recall*)

(resultado da relação entre o número de informações relevantes recuperadas e o total de informações relevantes armazenadas) foi de 98,03% e a de Precisão foi de 91,85%, para um total de 650 quadros de uma base de imagens que foi obtida de forma manual, a partir de duas partidas válidas pela *UEFA Champions League* 2003 e 2007.

O Mapa de Saliência pode ser considerada uma representação que enfatiza determinadas regiões da imagem. Esta foi a abordagem apresentada por Pei et al. (2009), como forma de detectar os prováveis candidatos a bola; em seguida, o mapa inicial é estendido para o mapa de saliência binária e, por fim, um filtro de *Kalman* é utilizado para detectar a posição da bola.

O pré-processamento, no trabalho de Pei et al. (2009), é iniciado a partir da detecção do campo do jogo, por meio do modelo cromático HSV; em seguida, o mapa de saliência é obtido levando-se em consideração três características primordiais: intensidade da imagem, contraste verde-vermelho e contraste azul-amarelo; em seguida, o mapa de saliência é convertido para nível de cinza, recebe a aplicação do filtro de *Wiener* e é convertido em imagem binária. Finalizando, alguns candidatos são eliminados por meio de parâmetros, como área do objeto e relação entre essa área e o mínimo quadrado capaz de contorná-la. A reprodução fiel do algoritmo não se torna possível por que a maneira como os mapas de saliência foram construídos não está explicada em um suficiente nível de detalhes. Com relação aos resultados apresentados, o número de vídeos utilizados nos experimentos não foi informado, o que leva a questionar se as taxas de Revocação e de Precisão, iguais a 98,9% e 95,9%, respectivamente, podem ser consideradas representativas.

De posse do local onde acontece a partida de futebol, o gramado, e o principal elemento, a bola, o próximo passo é a detecção dos sujeitos do jogo, os jogadores. Na Seção 2.3 será apresentada uma síntese das pesquisas existentes nesta área e também uma análise dos problemas e dificuldades em realizar tal tarefa. Na Tabela 2.1, sumarizam-se os trabalhos descritos nesta seção.

Tabela 2.1 – Sumário de técnicas revisadas para Detecção de Bola.

Trabalhos	Técnicas Utilizadas	Outras Características	Detecção Gramado	Resultados
D´ORAZIO ET. AL. (2004)	- Transformada Circular de Hough - Rede MLP - Wavelet	Detecção de semicírculos	NÃO	Taxa de Detecção 97%
YU ET. AL. (2007)	- Grafos Ponderados - Algoritmo de Viterbi	Informações de tamanho e forma	SIM	Resultados Objetivos Inexistentes
PALLAVI ET. AL. (2008)	- Transformada Circular de Hough - Filtragens de candidatos - Grafos Dirigidos - Programação Dinâmica	Técnicas diferentes, de acordo com o tipo de tomada	NÃO	Revocação = 98,03% Precisão = 91,85%
PEI ET. AL. (2009)	- Filtro de Kalman - Mapa de Saliência	Não Apresenta	SIM	Revocação = 97,3% Precisão = 98,3%

2.3. Detecção de Jogadores

A detecção e/ou rastreamento de jogadores é necessária na determinação de informações importantes de uma partida de futebol como, por exemplo, a velocidade e a distância percorrida pelos jogadores durante a partida e a região de maior movimentação no campo, dentre outras.

O trabalho apresentado por Figueroa, Leite e Barros (2006) utiliza a técnica de subtração de *background* para a segmentação de jogadores no campo de futebol. Primeiramente, é realizada a extração do *background* do campo, ou seja, o campo sem a presença de jogadores; em continuação, é realizada a diferença entre a imagem corrente, a qual se deseja segmentar, e o *background* extraído. A imagem diferença é então binarizada com um limiar

predefinido e um filtro morfológico é utilizado para a extração de ruídos. Por fim, os *pixels* conectados são rotulados e definidos como *blobs*. O processamento descrito é ilustrado na Figura 2.3:

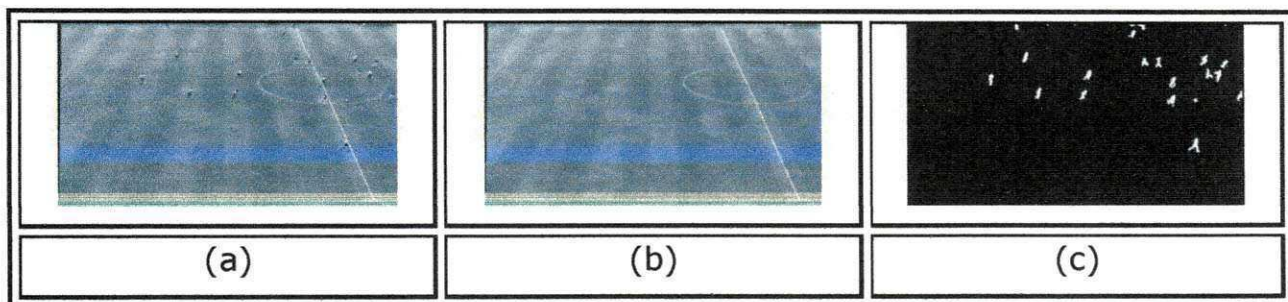


Figura 2.3 – Ilustração do método de Figueroa, Leite e Barros (2006): (a) Quadro de uma partida de futebol; (b) Background da imagem; (c) Blobs na imagem segmentada.

A abordagem utilizada no trabalho de Figueroa, Leite e Barros (2006) é muito simples, não sendo, entretanto muito aplicada na prática, devido à necessidade de obtenção do *background* para o processamento do algoritmo. Resultados numéricos não foram apresentados com relação à detecção de jogadores, uma vez que esta etapa é utilizada como pré-processamento para o rastreamento, propósito principal do artigo.

Spagnolo et al. (2007) apresentam a detecção de jogadores como uma etapa para a separação dos jogadores de cada time. Mediante a subtração da imagem desejada pelo *background*, o qual foi obtido de forma manual por intermédio da filmagem do gramado sem a presença de jogadores, são encontrados todos os jogadores presentes no campo. Separadamente, cada jogador tem seu histograma calculado e os que forem mais parecidos, ou seja apresentarem menos variações, são considerados com sendo do mesmo time. A abordagem apresentada é um pouco complexa por fazer uso de derivadas em conjunto com uma análise temporal, no tocante ao rastreamento de jogadores.

Com relação aos resultados, mais uma vez não foram apresentados pelo fato do trabalho principal ser a detecção de jogadores do mesmo time por intermédio do uso de histogramas de cores. Como o objetivo principal deste trabalho não era especificamente a detecção de jogadores, não é possível concluir que o método proposto é eficiente.

Beetz et al. (2007) apresentam o sistema ASPOGAMO, que tem a função de determinar as coordenadas e a trajetória dos jogadores durante uma partida de futebol. Para que seja realizada a detecção da trajetória deve-se, mais uma vez, realizar a detecção do jogador no campo, que nesse trabalho é feita por meio da segmentação dos jogadores em *blobs*, assim como no trabalho de Figueroa, Leite e Barros (2006). Após esta segmentação, os candidatos são analisados por meio de restrições de tamanho, modelos de cores e centro de massa, para que seja confirmado que o elemento encontrado realmente seja um jogador. A abordagem apresentada parece simples e apresenta resultados promissores: segundo os autores, o algoritmo proposto obteve taxas de detecção superiores a 90%. Uma melhoria que poderia aumentar os resultados seria a consideração de informação temporal, uma vez que apenas observações estáticas das imagens são consideradas no trabalho.

O trabalho desenvolvido por Arnaut e Costa (2009) apresenta uma técnica que segue uma abordagem probabilística, utilizando a variação da intensidade luminosa das imagens para realizar a detecção/rastreamento dos jogadores, em uma sequência de quadros. De Bem (2007) apresenta, em seu trabalho de mestrado, o rastreamento de alvos em sequências de imagens. Os referidos alvos devem possuir forma heterogênea e esse rastreamento se dá por intermédio do uso de dois filtros de *Kalman*, o primeiro para estimar a translação dos pontos e o segundo para estimar a velocidade de deslocamento que ocorre entre as imagens subsequentes. Arnaut e Costa (2009) utilizaram a técnica de De Bem (2007) para o caso específico da detecção de jogadores de futebol, conforme esquema apresentado na Figura 2.4.

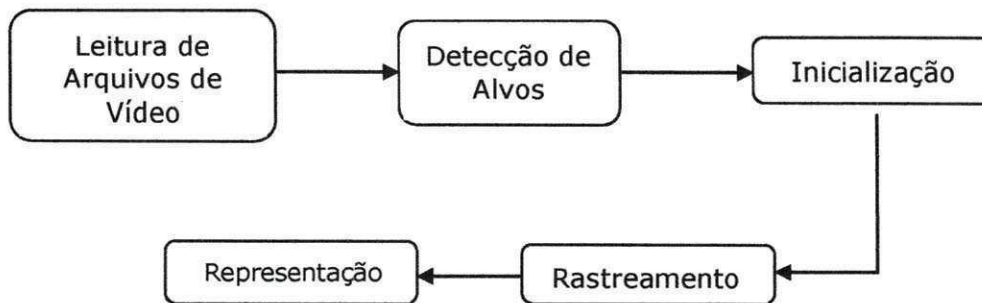


Figura 2.4 – Esquema para detecção de jogadores utilizando por Arnaut e Costa (2009).

Na detecção de alvos, no trabalho de Arnaut e Costa (2009), é utilizada a técnica de subtração de fundo, realizada *pixel a pixel*, a partir de uma imagem de referência do fundo da cena (*background*). O algoritmo ainda apresenta uma diferença em relação aos algoritmos apresentados anteriormente, ao utilizar contornos para os alvos segmentados construídos por *B-splines* (capaz de representar desde contornos triviais até formas complexas, por intermédio de um pequeno número de parâmetros), o que permite uma descrição mais precisa dos alvos quando comparada à técnica do mínimo retângulo (*Minimum Bounding Rectangle*) (BRINKHOFF; KRIEGEL; SCHNEIDER, 1993). A técnica de Arnaut e Costa (2009) parece promissora, levando-se em conta os resultados visuais apresentados, mas a ausência de uma análise estatística não permite confirmar esta afirmação.

No trabalho de Khatoonabadi e Rahmati (2009), uma etapa anterior à detecção de jogadores é a detecção da grama. A grama é detectada por meio de uma abordagem similar à utilizada por Ekin, Tekalp e Mehrotra (2003), com histogramas no espaço de cores HSI. Depois de realizada a detecção de grama um algoritmo de crescimento de regiões, em conjunto com a aplicação do filtro Mediana 3x3, é aplicado na região de grama, com a finalidade de remover possíveis ruídos, tal como as linhas do campo.

Khatoonabadi e Rahmati (2009) ainda definem algumas métricas de forma empírica, para eliminar prováveis elementos da imagem que não seriam jogadores, como a área do retângulo que envolve o elemento, que deve possuir valor maior que 50, e a razão entre as medidas dos lados, que

deve possuir valor máximo de 6. Um ponto positivo do trabalho é que existe uma análise comparativa com dois outros algoritmos: o algoritmo de Ekin, Tekalp e Mehrotra (2002) e o de Yoon, Bae e Yang (2002). Realizada a detecção de jogadores, calculam-se os resultados por intermédio das métricas *Predictivity* (razão entre TD (*True Detection*) e TD + FPD (*False Positive Detection*)) e *Sensitivity* (razão entre TD e TD + FND (*False Negative Detection*)). No trabalho de Khatoonabadi e Rahmati (2009), os resultados foram 89,79% para *Predictivity* e 86,97% para *Sensitivity*, enquanto que no trabalho de Ekin, Tekalp e Mehrotra (2002), os valores foram 75,32% e 74,82%, respectivamente. No trabalho de Yoon, Bae e Yang (2002) *Predictivity* ficou em 87,2% e *Sensitivity* em 81,07%, para base de imagens diferentes e rotuladas pelos autores. Alguns detalhes de implementação não foram bem descritos no artigo, como a aplicação do algoritmo de crescimento de regiões, dificultando a replicação do mesmo. Além disto, a obtenção da razão entre as medidas dos lados e do valor da área das figuras poderia ter sido recuperada de maneira mais confiável, por meio de um estudo estatístico.

A Tabela 2.2 contém uma sumarização das abordagens apresentadas. Finalizado o processo de detecção de grama, tomadas, bola e jogadores, resta ainda a detecção das linhas do gramado. Este tema será explorado na Seção 0, com a discussão dos trabalhos relevantes.

Tabela 2.2 – Sumário de técnicas revisadas para Detecção de Jogadores.

Trabalhos	Técnicas Utilizadas	Outras Características	Detecção Gramado	Resultados
FIGUEROA ET. AL. (2006)	- Subtração de Background Filtragem morfológica (abertura + fechamento) Grafos Orientados	- Uso de grafos e rastreamento de jogadores	NÃO	Não apresentam
SPAGNOLO ET. AL. (2007)	- Subtração de Background -Informações temporais	- Uso de informações de energia.	NÃO	DR = 95,26% FAR = 4,81%

BEETZ ET. AL. (2007)	- Segmentação em blobs - Modelos de cor aprendidos por marcação manual	- Restrições de forma e tamanho do jogador	SIM	Taxas de detecção superiores a 90%
KHATOONABAD I e RAHMATI (2009)	- Filtro da Mediana 3x3 - Operador de Abertura - Histogramas	- Restrições para eliminar falsos candidatos	SIM	Predictivity = 89,79% Sensitivity = 86,97%

2.4. Segmentação das Linhas do Campo

As linhas do campo são de fundamental importância em uma partida de futebol, por delimitarem a região em que os acontecimentos da partida são válidos. Sua detecção, em conjunto com a detecção da bola, pode ser utilizada para diminuir as dúvidas em certos lances, uma vez que a FIFA, entidade máxima do futebol mundial, está revendo o uso de tecnologias como forma de auxílio à decisão dos árbitros (VIVIANI, 2010).

No trabalho desenvolvido por Ekin, Tekalp e Mehrotra (2003) é apresentada uma abordagem para detecção das linhas paralelas da grande área. Esta detecção é importante pelo fato de que os principais eventos de uma partida de futebol acontecem no interior ou nas proximidades da grande área. A partir do momento em que é possível detectar os quadros que contêm a grande área, um vídeo poderá ser produzido, apenas com lances ocorridos naquela região. Naquele trabalho, os *pixels* das linhas desejadas serão aqueles com maior valor de gradiente, calculados usando uma máscara *Laplaciana 3x3*, apresentada na Equação 2.2. Por fim, a transformada de *Hough* é empregada para detectar as linhas paralelas.

$$h = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -8 & 1 \\ 1 & 1 & 1 \end{pmatrix} \quad (2.2)$$

Ekin, Tekalp e Mehrotra (2003) apresentam apenas a detecção de linhas paralelas, ou seja, as linhas da grande área do jogo, não detectando linhas laterais, que também são importantes em uma partida de futebol. Os resultados de Revocação em 83,7% e Precisão em 98% inspiram confiança na abordagem proposta; entretanto, a pouca quantidade de imagens envolvidas na obtenção dos resultados (61 imagens utilizadas para a detecção de linhas) e a insuficiência de detalhes de implementação são alguns dos pontos negativos da abordagem.

Cai e Tai (2005) propuseram um algoritmo para a detecção das linhas do campo que consiste em 3 etapas: extração da cor do gramado, remoção de quadros e objetos que não são desejados e por fim, a detecção das linhas do gramado. A extração da cor do gramado é obtida de forma simples, por intermédio do histograma de cores do modelo cromático RGB, o que ocasiona falhas na detecção devido à simplicidade daquele modelo. A segunda etapa do processamento consiste em remover os quadros do vídeo que não contenham as linhas do gramado, tais como propagandas e filmagens da torcida, dentre outros. Este processamento utiliza a cor dominante do quadro, isto é, o quadro analisado deverá ser eliminado caso contenha menos que 50% de cor dominante (verde do gramado).

A etapa final do algoritmo de Cai e Tai (2005) é subdividida em 4 subetapas: remoção de imagens em *zoom*, remoção de pixels não brancos (que representam elementos que não são linhas devendo então ser eliminados), aplicação do algoritmo conhecido como *line patching* (explicado mais adiante) e a remoção de objetos não desejados. Para a remoção de imagens com zoom em jogadores, foi realizado um estudo experimental que definiu um limiar igual a 34112 em relação à quantidade de *pixels* definidos como grama, para que a imagem seja considerada zoom (tomada de curta distância). Portanto, caso um quadro possua um número menor que esse de *pixels* de grama, deverá ser eliminado. A remoção dos *pixels* não brancos ocorre eliminando os que estiverem abaixo do limiar definido para os três espaços de cores de forma empírica: RGB = (200, 210, 190). O algoritmo

conhecido como *line patching*, é responsável por preencher as linhas que estiverem quebradas ou com algum ruído. Por fim, os elementos indesejados, como jogadores, serão eliminados por meio de algoritmos de vizinhança. O método apresentado é bastante simples e de fácil implementação; entretanto, não se pode ter confiança na sua eficácia, pois apenas imagens resultantes do algoritmo são apresentadas, inexistindo qualquer análise objetiva.

Lotfi e Pourreza (2006) desenvolveram uma abordagem que tem como principal objetivo a realização de uma sumarização automática de um vídeo de futebol. A estrutura utilizada no trabalho é mostrada na Figura 2.5.

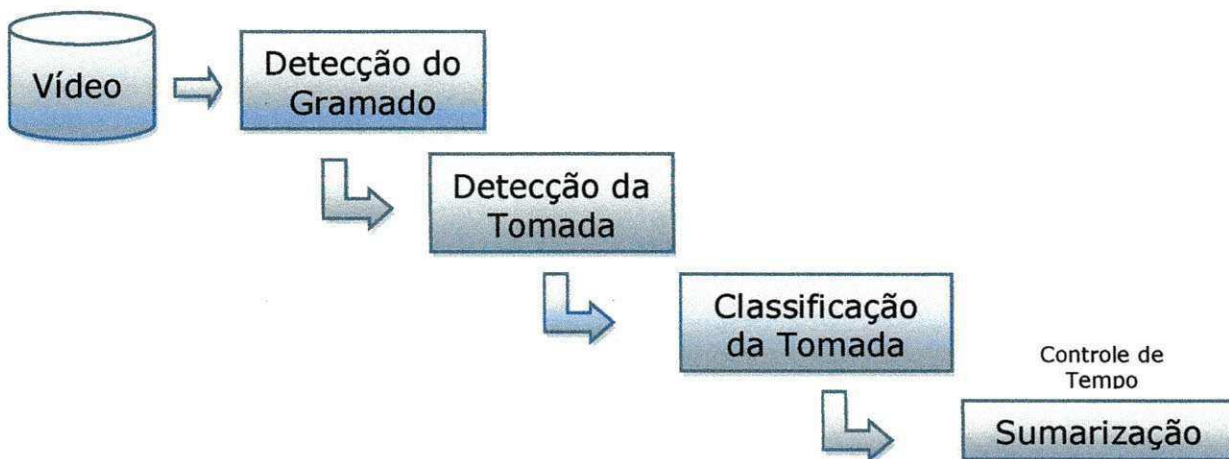


Figura 2.5 – Estrutura da abordagem proposta em Lotfi e Pourreza (2006).

Na detecção de grama, é utilizada por Lotfi e Pourreza (2006) a técnica de subtração entre histogramas no espaço HSI de imagens subsequentes, seguindo a abordagem apresentada em Ekin, Tekalp e Mehrotra (2003); em seguida, é realizada a classificação da tomada, dividida, de acordo com o autor, em 4 tipos, a saber: 1) meio campo, 2) ataque, 3) curta distância e 4) fora do campo. A tomada do tipo Fora de Campo é obtida na etapa 01, a partir do percentual de grama presente no quadro e, caso o quadro apresente menos que 40% dos *pixels* da imagem, ele será considerado tipo 4. Para distinguir as tomadas tipos 2 e 3, o autor utiliza a transformada de *Hough*, responsável pela detecção de linhas no campo, fazendo uso da angulação da linha que delimita o campo de jogo. Na Figura 2.6, apresenta-se um exemplo desta abordagem:

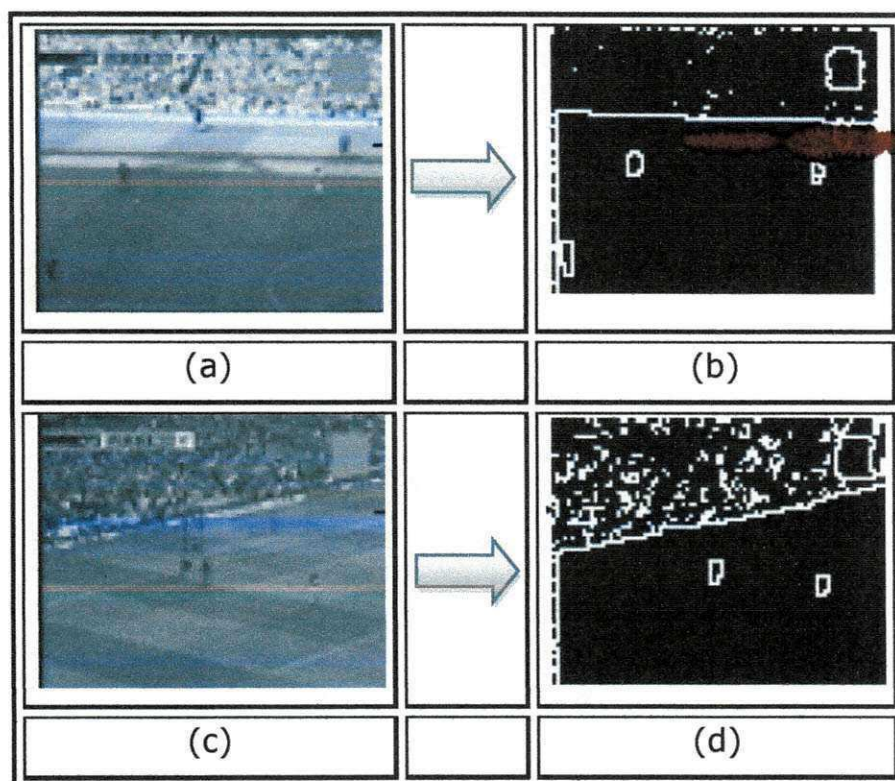


Figura 2.6 – Abordagem de Lofti e Pourreza (2006): (a) Tomada do tipo 1; (b) Quadro após processamento; (c) Tomada do tipo 2; (d) Quadro após processamento.

Conforme se pode verificar, o quadro do tipo 2 possui uma angulação maior em comparação com a do quadro do tipo 1, fato constatado empiricamente, o que valida o uso desta abordagem. Alguns resultados estatísticos apresentam uma acurácia de entre 81 e 92% na obtenção dessas tomadas, por meio do uso da detecção de linhas. A base é formada por 6 vídeos, sendo 1 da *Premier League* 2004 e 5 pequenos vídeos da Euro 2004, totalizando 5,5 horas de vídeos, base considerada relevante se comparada aos trabalhos descritos anteriormente.

Conforme observado no trabalho de Khatoonabadi e Rahmati (2009) e sabido por todos os que acompanham o futebol, as linhas do gramado são brancas. O problema é que outros elementos presentes em uma partida de futebol também podem ser brancos, tais como bola, logomarcas de emissoras e placas de publicidade, dentre outros, tornando a detecção das linhas do campo mais complexa que a detecção da grama, que também se torna complexa quando o uniforme dos jogadores possui a mesma cor (verde) do

gramado. Para solucionar este impasse, os autores propuseram a aplicação de um filtro bidimensional Laplaciano, h , definido pela equação (2.3), aplicado à imagem em níveis de cinza com a finalidade de tornar a imagem mais homogênea. Após a aplicação do filtro, a região de Não grama da imagem é removida por meio da aplicação de um algoritmo que elimina regiões 10x10 que contenham mais que 75% de grama. Por fim, são eliminados, da imagem, os *pixels* com baixos valores RGB e é aplicada a Transformada de *Hough* para selecionar os candidatos à linha do campo.

$$h = \begin{pmatrix} -2 & 1 & -2 \\ 1 & 4 & 1 \\ -2 & 1 & -2 \end{pmatrix} \quad (2.3)$$

Uma abordagem alternativa apresentada em Khatoonabadi e Rahmati (2009) é a detecção de linhas de campo a partir da informação temporal, ou seja, considerando-se vários quadros subsequentes do vídeo. Esta abordagem tem o propósito de reduzir o erro na detecção bem como diminuir o tempo de processamento, substituindo a Transformada de *Hough* pela técnica conhecida como *Extended Kalman Filter* (EKF) (MILLS; PRIDMORE; HILLS, 2003). Apesar de complexo, o trabalho de Khatoonabadi e Rahmati (2009) é rico em detalhes, o que facilita sua reprodução; além disto, são mostradas duas abordagens distintas para o problema de detecção de linhas. O objetivo principal é a detecção e o rastreamento de jogadores; sendo assim, uma lacuna presente é a inexistência de análise estatística para a detecção de linhas, isto é, os autores fornecem apenas uma avaliação subjetiva com bons resultados visuais, porém não é possível inferir objetivamente sobre o desempenho do detector de linhas.

Sun e Liu (2009) apresentaram um método para detecção de linhas que, como na maioria dos trabalhos discutidos anteriormente, tem início com a detecção da região de cor dominante, utilizando-se o espaço de cores HSI.

É apresentado o operador morfológico *White top-hat*, cuja função é realçar o contraste em ambientes não uniformes. O operador *White top-hat* é apresentado na Equação (2.4)

$$T_w(f) = f - f \circ b, \quad (2.4)$$

em que f é a imagem original, b a imagem em nível de cinza e \circ representa a operação morfológica de abertura. O algoritmo consiste, basicamente, em 3 passos:

1. Aplicação do operador da Equação (2.4) na imagem em nível de cinza;
2. Uso de máscaras predefinidas para eliminação de ruídos do campo;
3. Aplicação do operador morfológico *Thinning*, cujo objetivo é obter formas mínimas da imagem pelo processo de afinamento.

O resultado da detecção pode ser verificado na Figura 2.7

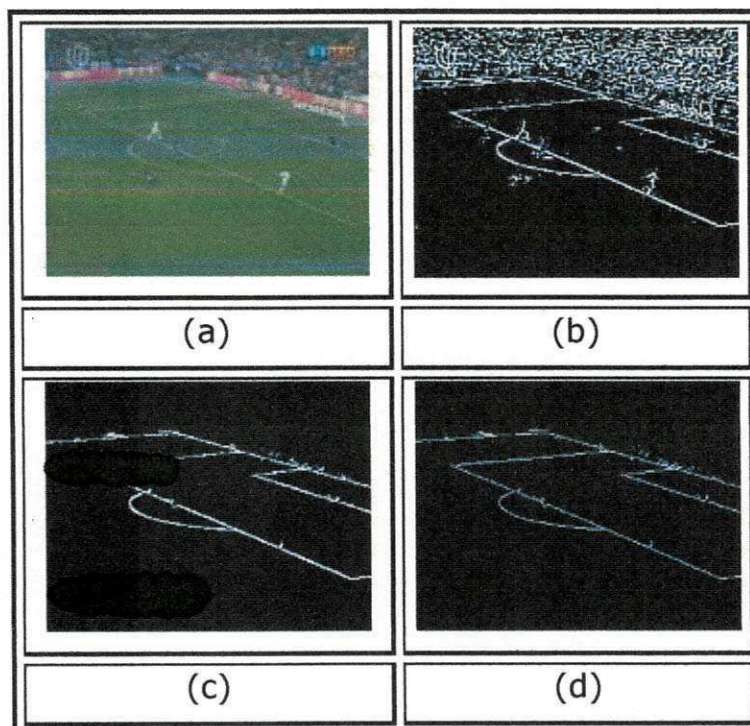


Figura 2.7 – Resultado do método de Sun e Liu (2009): (a) Quadro de uma partida de futebol; (b) Background da imagem; (c) Blobs na imagem segmentada; (d) Imagem após a aplicação do Thinning.

Ao contrário dos trabalhos revisados anteriormente, este apresenta uma avaliação experimental com as métricas de Revocação e Precisão. Para o experimento foram utilizados cerca de 200 quadros (não foi informada a

quantidade de vídeos) levando-se em consideração as linhas de meio-campo, linha de fundo e linhas da grande área. Os melhores resultados foram para a detecção de linhas de meio-campo, visto que apresentaram 95,4% de Revocação e 100% de Precisão. Na Tabela 2.3 sumarizam-se as técnicas apresentadas nesta seção.

Tabela 2.3 – Sumário de técnicas revisadas para Segmentação das Linhas do Campo.

Trabalhos	Técnicas Utilizadas	Outras Características	Deteção Gramado	Resultados
EKIN ET. AL. (2003)	Máscara 3x3 Laplaciana Transformada de Hough	Deteção das linhas da grande área	SIM	Revocação = 83,7% Precisão = 98%
CAI E TAI (2005)	Remoção de quadros Line Patching	Informações empíricas para a cor	SIM	Resultados Objetivos Inexistentes
KHATOONABADI E RAHMATI (2009)	Transformada de Hough Aplicação de filtros	Abordagens distintas: estática e temporal	SIM	Resultados Objetivos Inexistentes
SUN E LIU (2009)	Operador white top-hat Algoritmo Thinning	Aplicações de máscaras para eliminação de ruídos	SIM	Revocação = 95,4% Precisão = 100%

2.5. Considerações Finais do Capítulo

Diante do exposto neste capítulo, constata-se que existem muitos elementos relevantes em uma partida de futebol, tais como bola, jogador, linhas e tomadas que podem ser utilizados na análise de vídeos. Não foi verificada, porém, na grande maioria dos trabalhos analisados, uma análise estatística objetiva bem fundamentada.

O trabalho desenvolvido nesta dissertação pretende suprir esta carência, realizando uma análise objetiva e mais aprofundada de técnicas atinentes aos aspectos de Segmentação do Gramado e das Linhas do Campo, bem como de Classificação de Tomadas. Para tanto foi utilizada a curva ROC (PRATI; BATISTA; MONARD, 2008).

O próximo capítulo contém a descrição das características das bases de imagens, tais como: horário e local dos jogos, quantidade de quadros de cada partida, tipo de quadro de cada partida, formato do vídeo e taxa de amostragem.

Capítulo 3

Base de Imagens

3.1. Segmentação da Grama e das Linhas

Para a realização dos testes relacionados à detecção de grama, a criação de uma base de imagens foi necessária com fins à validação da abordagem proposta e comparação com as demais selecionadas. A base de imagens é composta de 300 imagens, obtidas de forma aleatória, a partir de uma coleção de 15 vídeos (20 quadros por vídeo), dos quais 11 foram da última Copa do Mundo de Futebol, realizada na África do Sul, no ano de 2010. Os outros 4 vídeos são de diferentes campeonatos, como a Copa dos Campeões da UEFA e a *Champions League*, disputadas no continente Europeu, o Campeonato Paulista e a Copa do Brasil.

O primeiro vídeo foi capturado da partida realizada entre as seleções da Alemanha e Austrália, em Durban no horário local das 15h30. O vídeo possui o formato *AVI*, uma taxa de dados de 96kbps e uma taxa de amostragem de 30 quadros por segundo. Com relação aos 20 quadros escolhidos para compor a base, 11 foram de longa distância, 9 de média distância e 1 de curta distância, totalizando 63,39% de *pixels* grama e 36,61% de *pixels* não grama.

A partida entre Argentina e Alemanha, realizada na Cidade do Cabo, no horário das 11h00, compõe o segundo vídeo da base de imagens, também no formato *AVI* e com uma taxa de amostragem de 30 quadros por segundo, porém com uma taxa de dados de 112kbps. As imagens obtidas estão assim

dispostas: 12 de longa distância, 4 de média distância e 3 de curta distância, em que 56,82% dos *pixels* foram considerados grama e 43,18% foram rotulados como não grama.

O terceiro vídeo que compõe a base corresponde a uma partida de futebol realizada na cidade de *Bloemfontein*, no horário das 11h00, entre as seleções da Argentina e Nigéria. O vídeo compartilha das mesmas configurações do primeiro vídeo, codificado no formato AVI, com uma taxa de dados de 96kbps e taxa de amostragem de 30 quadros por segundo. A taxa de *pixels* considerados grama foi de 58,46%, enquanto 41,54% dos *pixels* foram rotulados como não grama em 13 imagens de longa distância, 2 de média distância e 5 de curta distância.

A quarta partida marcou a final da *Champions League* do ano de 2009, realizada no Estádio Olímpico, em Roma. O vídeo em formato AVI possui taxa de dados de 47kbps e taxa de amostragem de 29 quadros por segundo com 11 quadros de longa distância, 4 de média distância e 5 de curta distância. *Pixels* considerados grama totalizaram 57,06% e *pixels* não grama foram 42,93%.

A partida entre Brasil e Holanda, válida pelas quartas de final da Copa do Mundo da África do Sul, foi utilizada como quinto vídeo da base criada. Em formato AVI, com uma taxa de dados de 34kbps, taxa de amostragem de 25 quadros por segundo, o vídeo é composto de 11 quadros de longa distância, 5 de média distância e 4 de curta distância. Com relação ao percentual de *pixels*, 63,82% são compostos de grama e 36,18% de não grama.

Ainda da Copa do Mundo 2010, foi escolhida outra partida realizada pela Seleção Brasileira, desta vez contra a equipe de Portugal, em partida ainda válida pela fase de grupos daquele mundial. O vídeo se encontra em formato AVI com uma taxa de 112kbps e taxa de amostragem de 30 quadros por segundo. A distribuição dos *pixels* foi composta de 72,09% de grama e 27,81% de não grama, pelo fato da grande quantidade de imagens de longa distância selecionadas, de 17, contra 2 de média distância e 1 de curta distância.

O sétimo jogo selecionado foi realizado entre Brasil e Chile, pelas oitavas de final do mundial da África do Sul. Em formato AVI, taxa de dados de 769kbps e taxa de amostragem de 30 quadros por segundo, este vídeo possui, nos quadros selecionados, 56,95% de *pixels* rotulados como grama e 43,05% rotulados como não grama. Este equilíbrio se deu por conta do grande número de tomadas de curta e média distância, que foram 5 e 6, respectivamente, contra 9 tomadas de longa distância.

O oitavo jogo a compor a base foi realizado na Cidade do Cabo, às 20:30h, entre as seleções de Camarões e Holanda, em partida válida pela Copa do Mundo da África 2010. Em formato AVI, com uma taxa de dados de 112kbps e uma taxa de amostragem de 30 quadros por segundo, possui 13 imagens de longa distância e 7 de média distância, não apresentando nenhuma de curta distância. O percentual de *pixels* ficou equilibrado com 56,54% de grama e 43,46% de não grama.

O nono jogo utilizado foi realizado entre Espanha, que se sagraria campeã do mundial da África do Sul e Portugal, nas oitavas de final daquele mundial. Em formato AVI, taxa de dados 115kbps e taxa de amostragem de 25 quadros por segundo, os quadros selecionados obedeceram à seguinte disposição: 3 imagens de curta distância, 4 de média distância e 13 de longa distância. *Pixels* rotulados como grama compõem 55,65% do total enquanto 45,35% foram rotulados como não grama.

O vídeo da partida realizada entre França e México foi selecionado como sendo o décimo vídeo que forneceu quadros para compor a base, principalmente pelo fato da camisa da seleção mexicana ser verde, fato interessante para teste do algoritmo de detecção da grama. O vídeo em formato AVI, taxa de dados de 39kbps e taxa de amostragem de 25 quadros por segundo, possui 67,52% de *pixels* grama contra 32,48% de não grama, distribuídos em 10 imagens de longa distância, 8 de curta distância e 2 de média distância.

O vídeo de número 11 da base é composto pela partida entre Corinthians-SP e Internacional-RS, pela final da Copa do Brasil de 2009, realizada na cidade de Porto Alegre. Em formato *AVI*, o vídeo possui uma taxa de dados de 1131kbps e taxa de amostragem de 29 quadros por segundo. Com 12 imagens de longa distância, 5 de média distância e 3 de curta distância, este foi o primeiro vídeo em que o percentual de *pixels* grama foi inferior ao percentual não grama, 46,26% e 53,73%, respectivamente.

O décimo segundo vídeo da base foi capturado da partida realizada entre Manchester United e Chelsea pela final da *Champions League* 2008, na cidade de Moscou, na Rússia e possui formato *AVI*, taxa de dados de 89kbps e taxa de amostragem de 25 quadros por segundo; as 20 imagens selecionadas estão distribuídas da seguinte forma: 9 imagens de longa distância, 4 de média distância e 7 de curta distância, totalizando 60,63% de *pixels* rotulados como grama e 39,36% como não grama.

O vídeo com os melhores lances da partida disputada entre Palmeiras e Americana, pelo Campeonato Paulista de 2010, em formato *AVI*, taxa de dados de 1211kbps e taxa de quadros de 29 por segundo, foi definido como décimo terceiro vídeo da base. Os quadros de longa distância totalizaram 11, os de média distância totalizaram 6, e os de curta distância 2, além de 1 quadro de filmagem da torcida, totalizando 74,09% de grama e 25,91% de não grama.

O décimo quarto vídeo, em formato *AVI*, com taxa de dados de 50kbps e taxa de amostragem de 29 quadros por segundo, foi obtido do jogo entre Paraguai e Espanha, pelas quartas de final da Copa do Mundo 2010. Em partida muito disputada, a Espanha venceu por 1x0. Foram selecionadas, também de forma aleatória, 20 imagens, em que 15 imagens são de longa distância, 1 de média distância, 3 de curta distância, além de 1 imagem com logo da FIFA. Com relação aos percentuais de *pixels*, 61,69% foram rotulados como grama e 38,31% como não grama.

Para finalizar a base de imagens, foi utilizada também uma partida válida pela Copa do Mundo de 2010, entre as seleções de Uruguai e Holanda, única em formato *mp4* que não disponibiliza maiores informações acerca das taxas de amostragem e de dados. A partida, realizada na cidade de Porto Elizabeth, no horário local das 11h, terminou com a vitória da seleção Holandesa por 3x2. Dos quadros selecionados, 11 foram de longa distância, 6 de média distância e 3 de curta distância. O equilíbrio entre as imagens de longa e curta distância foi refletido na quantidade de *pixels* rotulados como grama e não grama, 56,11% e 43,89% respectivamente. Esta e outras informações podem ser verificadas na Tabela 3.2, que contém um sumário dos dados apresentados no presente capítulo.

Depois de construída a base para a segmentação do gramado, foi necessária uma adaptação para que fosse possível testar a segmentação das linhas, ou seja, as imagens de longa distância (184 imagens) da base marcadas como grama e não grama foram adaptadas para: Grama, Bola, Jogadores, Linhas e Outros (Torcida, Trave, Logo da Emissora, etc.). A Tabela 3.1 contém um sumário de como foi feita a relação entre o elemento rotulado e a respectiva cor selecionada para marcação.

Tabela 3.1 – Relação entre elementos e cores.

Elemento	Cor	R,G,B [0;255]
Grama	Preta	0,0,0
Jogador	Vermelha	255,0,0
Linhas	Branca	255,255,255
Bola	Azul	0,0,255
Outras	Verde	0,255,0

A rotulagem foi realizada desta forma para facilitar os cálculos realizados no processamento. Uma vez que a distância euclidiana utiliza potenciação e raiz quadrada, as cores primitivas facilitam o seu cálculo. Também o foi para que a base criada pudesse atender às necessidades de

desenvolvimentos futuros. Os quadros originais utilizados para a construção da base de imagens, a base de imagens rotulada para a segmentação do gramado e segmentação das linhas do campo, está disponível em: <http://lvc.computacao.ufcg.edu.br/bases-de-dados>.

3.2. Classificação de Tomadas

Para que a eficácia da Classificação de Tomadas fosse verificada, tornou-se necessária a construção de uma base de imagens rotuladas em longa distância e outras. Dos vídeos utilizados nesta base, três foram reaproveitados da base definida na Seção 3.1 (Alemanha x Austrália, Argentina x Alemanha e Espanha x Portugal) e outro novo vídeo foi incluído (Santos x Flamengo). A base é composta de 5391 imagens (quadros de vídeo) rotuladas como longa distância e 2974 rotuladas como outras, totalizando 8365 imagens. A disposição dos tipos de tomada por cada jogo está descrita na Tabela 3.1 e os quadros separados em duas classes: longa distância e outras, utilizados na avaliação experimental, estão disponíveis em: <http://lvc.computacao.ufcg.edu.br/bases-de-dados>.

Tabela 3.3 – Características da Base de Imagens criada para Classificação de tomadas.

Jogo/Características	longa distância	outras
Alemanha Vs Austrália	2100	0716
Argentina Vs Alemanha	1200	1050
Espanha Vs Portugal	1058	0685
Santos Vs Flamengo	1033	0523
TOTAL	5391	2974

3.3. Considerações Finais do Capítulo

Este capítulo apresentou como foram criadas as três bases de imagens necessárias para realizar a avaliação experimental e o próximo apresentará as abordagens propostas neste trabalho bem como detalhes da implementação realizada.

Tabela 3.2 – Características da Base de Imagens (Segmentação das Linhas).

Jogo/Características	Formato	Taxa de Dados (kbps)	Taxa de Amostragem (fps)	Ano	Horário	Evento	Rotulagem	
							grama (%)	não grama (%)
Alemanha Vs Austrália	AVI	96	30	2010	15h30min	Copa do Mundo	63,39	36,61
Argentina Vs Alemanha	AVI	112	30	2010	16h45min	Copa do Mundo	56,82	43,18
Argentina Vs Nigéria	AVI	96	30	2010	14h00min	Copa do Mundo	58,46	41,54
Barcelona Vs Manchester	AVI	47	29	2011	15h45min	Liga dos Campeões	57,06	42,93
Brasil Vs Holanda	AVI	34	25	2010	11h00min	Copa do Mundo	63,82	36,18
Brasil Vs Portugal	AVI	112	30	2010	11h00min	Copa do Mundo	72,09	27,81
Brasil Vs Chile	AVI	769	30	2010	15h30min	Copa do Mundo	56,95	43,05
Camarões Vs Holanda	AVI	112	30	2010	15h30min	Copa do Mundo	56,54	43,46
Espanha Vs Portugal	AVI	115	25	2010	15h30min	Copa do Mundo	55,65	45,35
França Vs México	AVI	39	25	2010	15h30min	Copa do Mundo	67,52	32,48
Corinthians Vs Internacional	AVI	1131	29	2009	21h50min	Copa do Brasil	46,26	53,73
Manchester Vs Chelsea	AVI	89	25	2007	15h45min	Liga dos Campeões	60,63	39,36
Palmeiras Vs Americana	AVI	1211	29	2010	16h00min	Camp. Paulista	74,09	25,91
Paraguai Vs Espanha	AVI	50	29	2010	15h30min	Copa do Mundo	61,69	38,31
Uruguai Vs Holanda	Mp4	-	-	2010	15h30min	Copa do Mundo	56,11	43,89

Capítulo 4

Abordagens Propostas

Ao iniciar o estudo sobre os aspectos relevantes associados a uma partida de futebol, observou-se uma sequência lógica na ordem do desenvolvimento das técnicas. Sendo assim, ficou definido, neste trabalho, que os três aspectos estudados, em ordem, seriam: Segmentação do Gramado, Segmentação das Linhas do Campo e Classificação de Tomadas.

A Segmentação do Gramado é de fundamental importância, pois define a área efetiva em que o jogo ocorre, reduzindo o espaço de busca do segmentador de linhas. Outro ponto para que este aspecto fosse o primeiro a ser pesquisado é sua influência na Classificação de tomadas, uma vez que Tomadas de longa distância geralmente possuem grande percentual de grama. A Segmentação das Linhas vem em seguida, por conta de ser, também, outro fator que pode auxiliar na Classificação de tomadas, também um dos objetivos desta pesquisa.

4.1. Segmentação do Gramado

Mediante o estudo de alguns vídeos, Firmino Jr. e Gomes (2007) definiram uma abordagem de classificação baseada nos k-vizinhos mais próximos. Um conjunto de 30 imagens foi obtido a partir da seleção de 6 imagens de cada um dos 5 vídeos de partidas de futebol, em diferentes horários e condições climáticas. Em cada imagem, foram selecionados alguns pontos (ou *pixels*, representado no espaço de cores RGB) pertencentes ao gramado e outros não pertencentes. Esses pontos foram utilizados como base para o treinamento do sistema ora apresentado, sendo os *pixels* das imagens, portanto, rotulados em dois grupos, grama e não grama.

O algoritmo consiste em computar a distância Euclidiana, segundo a Equação (4.1), entre o *pixel* da imagem analisada e um conjunto de “*k*” valores de *pixels* selecionados aleatoriamente da base de dados. A partir do produto entre as menores distâncias *pmd* - Equação (4.2) - de cada grupo (grama e não grama), obtém-se um valor em tom de cinza (pertencente ao intervalo 0-255) que representa uma estimativa da probabilidade do *pixel* ser grama.

$$pmd = \frac{((255 - MNVG) * MGv)}{255} \quad (4.1)$$

$$Dist(i_1, i_2) = \sqrt{(r_1 - r_2)^2 + (g_1 - g_2)^2 + (b_1 - b_2)^2} \quad (4.2)$$

em que:

- *MNVG* é a **Menor Valor** de distância entre o *pixel* e os elementos do conjunto **Não Grama**;
- *MVG* é a **Menor Valor** de distância entre o *pixel* e os elementos do conjunto **Grama**;
- i_n representa o *n*-ésimo *pixel* analisado;
- r_n, g_n, b_n representam os Componentes Vermelho (*Red*), Verde (*Green*) e Azul (*Blue*), respectivamente, do *n*-ésimo *pixel*.

Depois de realizado o mapeamento para tons de cinza, conforme a Equação (4.2), na qual, quanto menor o valor do *pixel*, maior a probabilidade do mesmo ser grama, é aplicado o filtro BLUR 9x9, que tem a função de suavizar as imagens e, por fim, o algoritmo realizará uma binarização simples, a partir de um limiar *I* obtido de forma experimental, conforme ilustrado na Figura 2.8. No trabalho originalmente desenvolvido por Firmino Jr. e Gomes (2007), é apresentada uma validação estatística não muito aprofundada dos resultados, uma vez que naquele trabalho comparam-se

apenas os percentuais de Erro Tipo I (o algoritmo marcou como grama e era não grama) e de Erro Tipo II (o algoritmo marcou como não grama e era grama) com os resultados apresentados no trabalho de Wang et al. (2004). O algoritmo de Firmino Jr. e Gomes (2007) obteve Erro Tipo I igual a 2,208% e Erro Tipo II igual a 6,684%, enquanto Wang et al. (2004) obtiveram 2,806% e 38,228%, respectivamente. A segmentação do gramado proposta por Firmino Jr. e Gomes (2007) obtinha os valores para os parâmetros " k " – Número de imagens escolhidas de forma aleatória e " l " – Limiar utilizado para a binarização de forma empírica, não percorrendo todas as possibilidades existentes para tais parâmetros. Nesta dissertação, aquele trabalho foi aperfeiçoado, pois todos os possíveis valores para estes dois parâmetros foram testados por meio de uma variação sistemática e analisados por intermédio da curva ROC.

Uma base de imagens (descrita no Capítulo 3) maior e mais diversificada foi criada nesta dissertação, com a finalidade de demonstrar que o algoritmo de Segmentação do Gramado desenvolvido apresenta uma taxa de segmentação melhor do que os demais trabalhos revisados. Desta alteração, dois arquivos de dados foram gerados, contendo 300 quadros de vídeo com seus pixels rotulados como grama ou não grama.

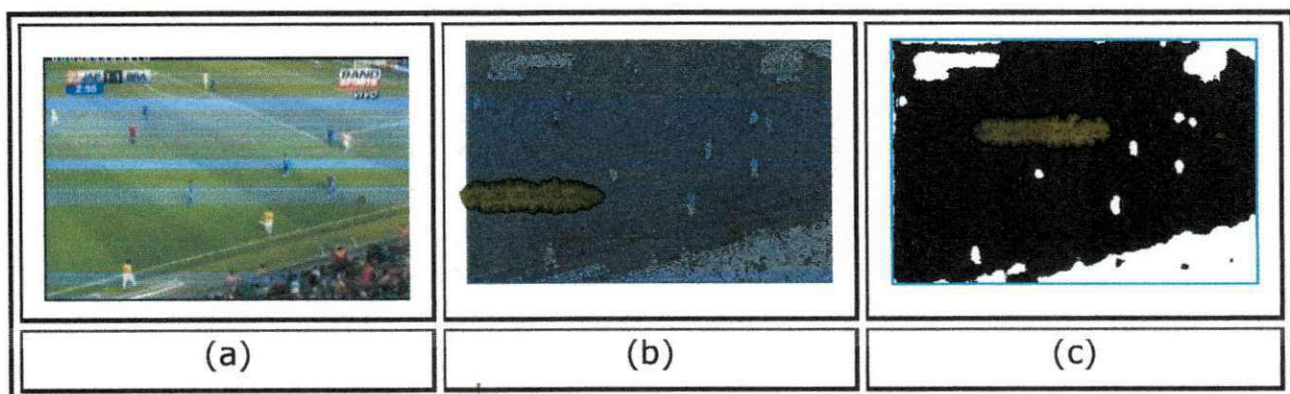


Figura 2.8 – Resultado da etapa de Segmentação do Gramado: (a) Imagem original; (b) Imagem que estima a probabilidade do pixel não ser Grama; (c) Imagem segmentada após aplicação do limiar.

4.2. Segmentação das Linhas do Campo

A Segmentação das Linhas surge como o segundo desenvolvimento da presente pesquisa. Após delimitar a região do gramado (Seção 4.1), deve-se buscar as linhas apenas na região delimitada como grama. Durante a revisão bibliográfica realizada, constatou-se que a transformada de *Hough* é a principal técnica utilizada na detecção de linhas em partidas de futebol. Porém, muitos dos trabalhos presentes na literatura não são suficientemente claros no que diz respeito ao pré-processamento utilizado na detecção, aos parâmetros utilizados e aos problemas que este tipo de transformada pode causar na detecção das linhas.

Algumas formas existentes podem ser facilmente parametrizadas, tais como círculos e linhas. Sendo assim, *Paul Hough* desenvolveu em 1962 a transformada que levou seu sobrenome e foi patenteada pela IBM. Esta transformada consiste basicamente em definir um mapeamento entre a imagem e os parâmetros necessários para definir a forma desejada. Após este mapeamento, as células são incrementadas e indicarão no final do processo, por meio da máxima local de um acumulador, quais os parâmetros correspondentes a forma especificada (ATIQUZZAMAN, 1992).

Algumas condições devem ser atendidas para que a transformada de *Hough* seja aplicada com sucesso à imagem, dentre elas a detecção de bordas e limiarização. Existem, na literatura, diversas formas de se realizar a detecção de bordas: Esqueletização, Filtro de *Sobel*, Filtro de *Canny*, bem como a limiarização, que pode ser adaptativa ou global simples, pelo método de *Otsu*, dentre outras (GONZALEZ; WOODS, 2008).

Nesta dissertação, foi primeiramente usada para realizar a detecção de bordas uma abordagem similar à Segmentação de Grama de Firmino Jr. e Gomes (2007). Porém, conforme se observa na Figura 2.9, os resultados não foram animadores, por conta da espessura das linhas e da influência do verde do gramado em sua cor, ocasionando falhas no processo de detecção.

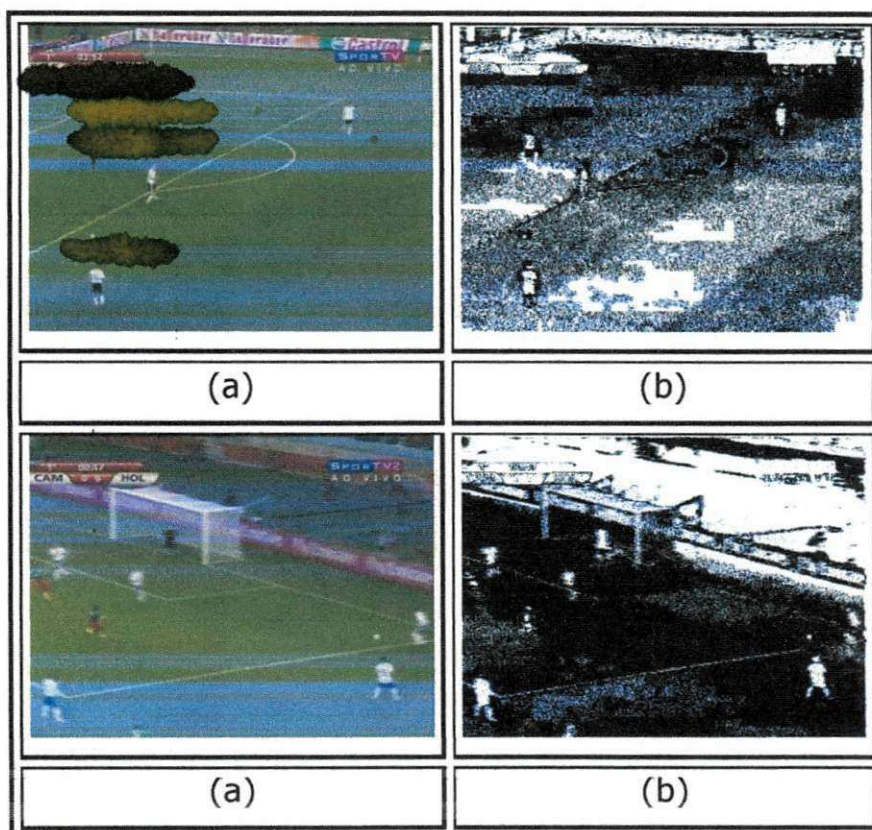


Figura 2.9 – Tentativa de segmentação das Linhas utilizando abordagem similar ao método de Firmino Jr. e Gomes (2007): (a) Imagem Original; (b) Imagem resultante.

Como esta primeira abordagem não foi satisfatória, foram testadas outras abordagens para a detecção de bordas, tais como *Canny*, *Sobel* e Esqueletização. No que diz respeito à limiarização, duas técnicas foram implementadas: adaptativa e global simples. A esqueletização consiste em se reduzir a figura analisada a uma cadeia de pixels, com a largura de apenas um pixel, levando-se em consideração a vizinhança e a variação claro-escuro da imagem (GONZALEZ; WOODS, 2008).

A esqueletização utilizada no desenvolvimento do detector de linhas foi baseada na aplicação de máscaras com diversas orientações, desenvolvida pela Universidade de *Edinburgh* (FISHER ET AL., 2003). A esqueletização consiste em utilizar a combinação das imagens geradas pela aplicação dos filtros M1 (responsável por detectar linhas horizontais), M2 (responsável por detectar as linhas verticais), M3 (responsável por detectar as linhas diagonais à esquerda) e M4 (responsável por detectar as linhas diagonais à direita), exibidos, respectivamente, nas Equações 4.3, 4.4, 4.5 e

4.6. Em seguida, deve-se efetuar uma normalização para lidar com valores negativos que foram gerados e, por fim, realizar a binarização para que a imagem possa ter suas linhas detectadas pela transformada de *Hough*.

$$M1 = \begin{pmatrix} -2 & -2 & -2 & -2 & -2 \\ 1 & 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 & 2 \\ 1 & 1 & 1 & 1 & 1 \\ -2 & -2 & -2 & -2 & -2 \end{pmatrix} \quad (4.3)$$

$$M2 = \begin{pmatrix} -2 & 1 & 2 & 1 & -2 \\ -2 & 1 & 2 & 1 & -2 \\ -2 & 1 & 2 & 1 & -2 \\ -2 & 1 & 2 & 1 & -2 \\ -2 & 1 & 2 & 1 & -2 \end{pmatrix} \quad (4.4)$$

$$M3 = \begin{pmatrix} 2,0 & 0,5 & -0,5 & -1,5 & -2,5 \\ 0,5 & 2,0 & 0,5 & -0,5 & -1,5 \\ -0,5 & 0,5 & 2,0 & 0,5 & -0,5 \\ -1,5 & -0,5 & -0,5 & 2,0 & 0,5 \\ -2,5 & -1,5 & -0,5 & -0,5 & 2,0 \end{pmatrix} \quad (4.5)$$

$$M4 = \begin{pmatrix} -2,5 & -1,5 & -0,5 & -0,5 & 2,0 \\ -1,5 & -0,5 & -0,5 & 2,0 & 0,5 \\ -0,5 & 0,5 & 2,0 & 0,5 & -0,5 \\ 0,5 & 2,0 & 0,5 & -0,5 & -1,5 \\ 2,0 & 0,5 & -0,5 & -1,5 & -2,5 \end{pmatrix} \quad (4.6)$$

Para que a binarização preserve as linhas do campo da melhor forma possível, foram analisados os histogramas de cinza de diversas partidas presentes na base. De acordo com a Figura 2.10, as imagens utilizadas apresentam histogramas similares após a normalização.

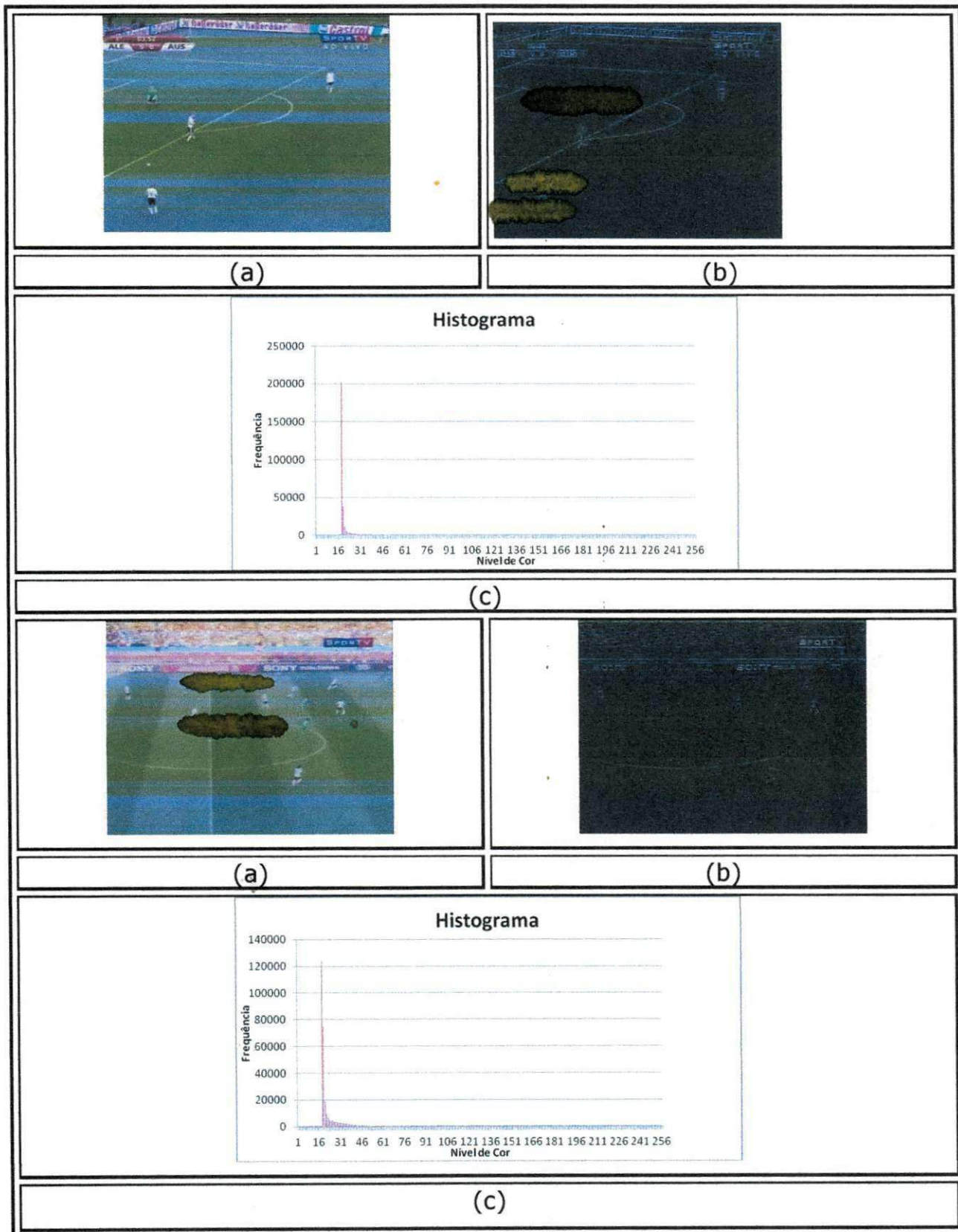


Figura 2.10 – Aplicação das máscaras para esqueletização: (a) Imagem Original; (b) Imagem resultante; (c) Histogramas da Imagem resultante.

Com a realização da análise visual dos histogramas, tornou-se evidente que o melhor limiar "**LB**" a ser utilizado na binarização da imagem, é fator do pico do histograma e deverá ser definido por meio da Curva ROC (ver Seção 5.2). Exemplos das imagens binarizadas que serviram de entrada para a detecção de linhas podem ser visualizados na Figura 2.11

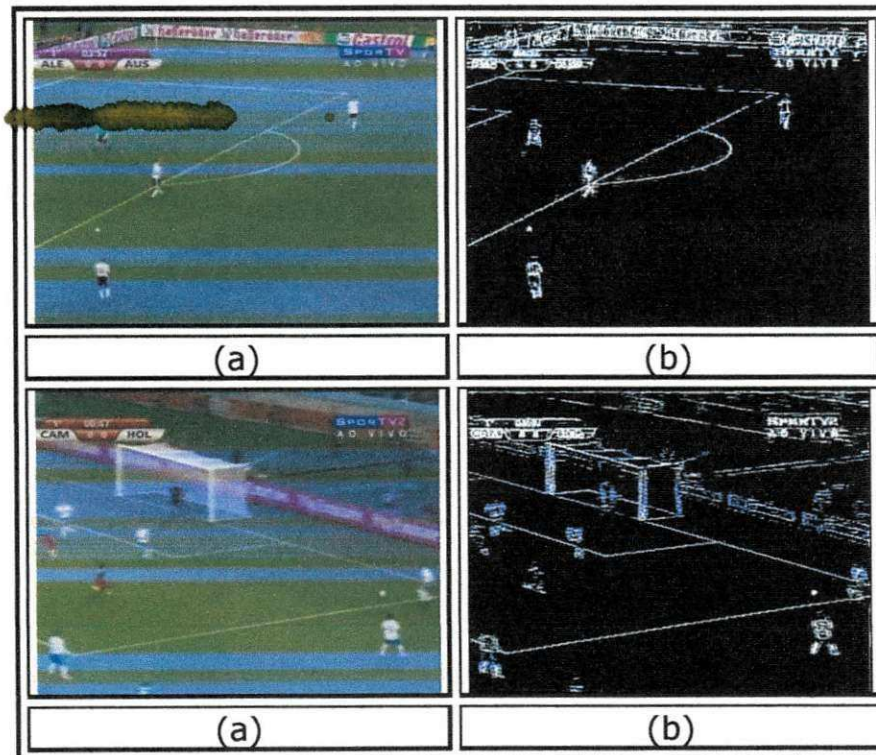


Figura 2.11 – Processo de Binarização: (a) Imagem original; (b) Imagem resultante.

Na Figura 2.12, observa-se que a transformada de *Hough*, aplicada a imagens de partidas de futebol, pode gerar problemas, por conta da espessura das linhas da grande área. Algumas linhas do campo são marcadas mais de uma vez, tornando-se necessária a implementação de um algoritmo de pós-processamento capaz de reduzir as linhas redundantes.

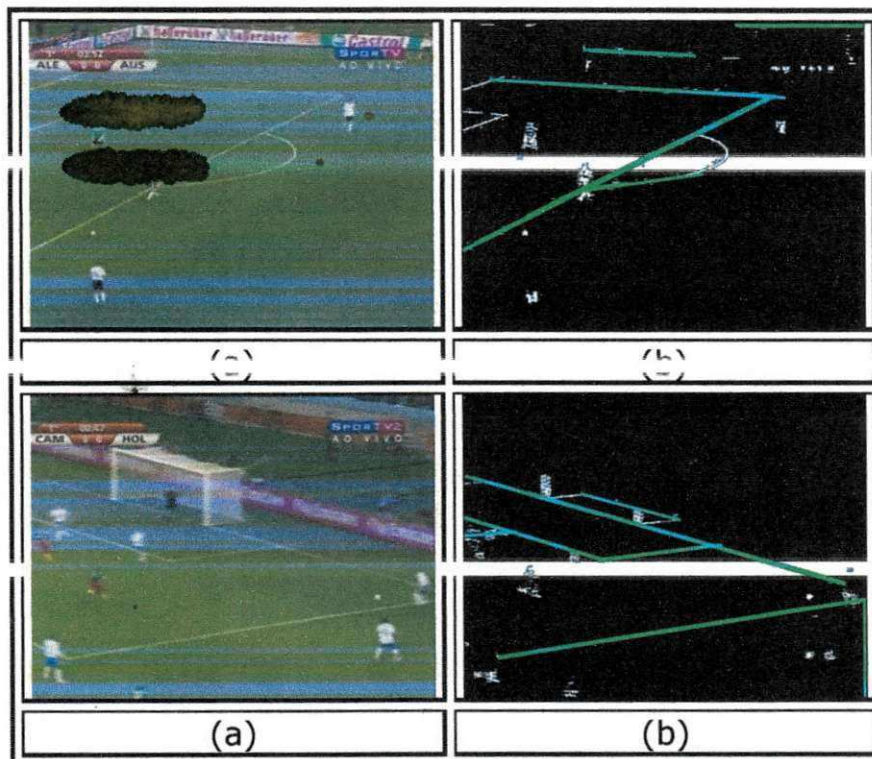


Figura 2.12 –Segmentação de Linhas com falhas: (a) Imagem original; (b) Imagem resultante.

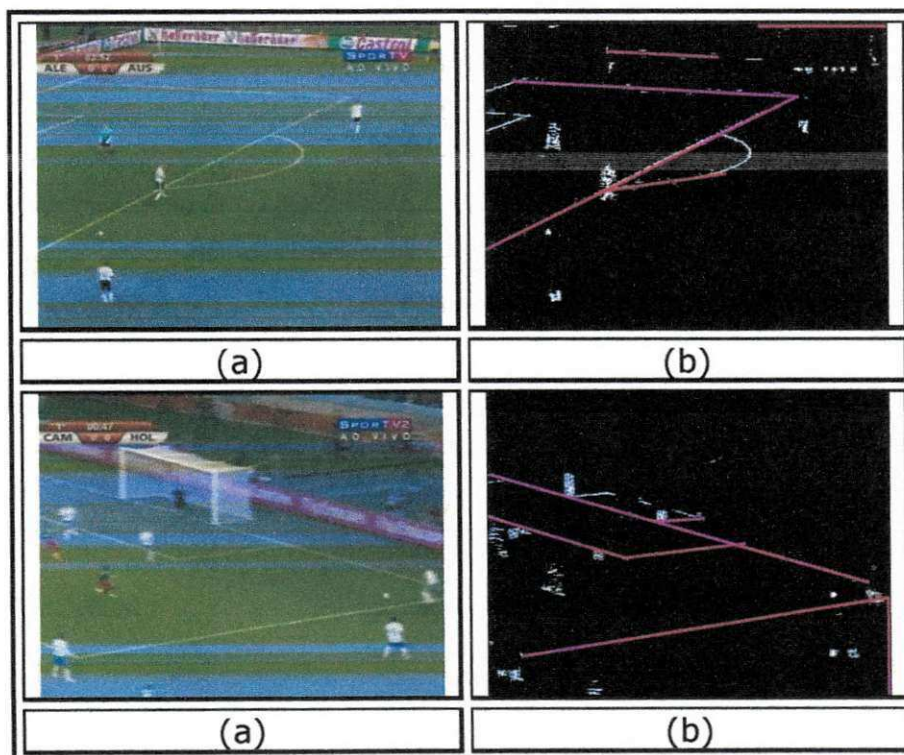


Figura 2.13 – Eliminação das linhas redundantes: (a) Imagem Original; (b) Imagem resultante.

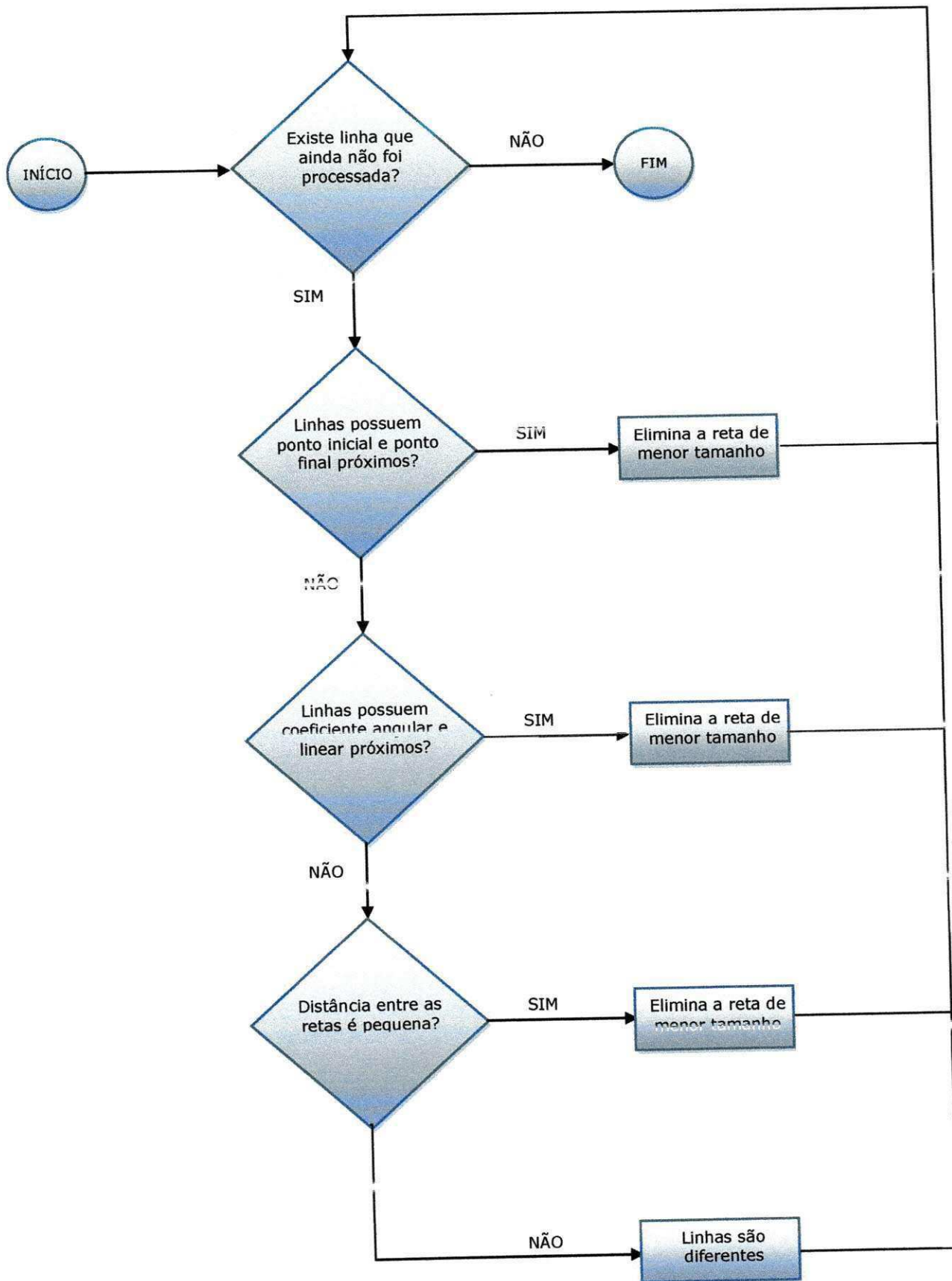


Figura 2.14 - Algoritmo de Eliminação de Linhas Redundantes.

Para que fosse possível realizar a redução das linhas, diversos atributos foram considerados, tais como coeficiente angular, coeficiente linear, ponto inicial, ponto final, ponto médio e tamanho das retas provenientes da Transformada de *Hough*. No diagrama exibido na Figura 2.14, resume-se o fluxo seguido para a eliminação das linhas redundantes, cujo resultado visual pode ser verificado na Figura 2.13.

Por fim, a aplicação desenvolvida salva as informações de cada imagem em um arquivo do tipo texto, o qual contém todas as informações relevantes às linhas recuperadas, tais como ponto inicial, ponto final e tamanho das linhas, podendo ser utilizado no cálculo de estatísticas da partida, sendo mais uma importante contribuição desenvolvida nesta dissertação.

4.3. Classificação de Tomadas

As tomadas de longa distância, caracterizadas por uma grande quantidade de jogadores, grande quantidade de grama e elementos com tamanhos muito reduzidos, em geral contêm os eventos mais significativos em uma partida de futebol, conforme anteriormente citado, durante a revisão bibliográfica. Sendo assim, sua identificação é de importância fundamental, caso se deseje um vídeo contendo os lances mais significativos. Neste trabalho, a classificação foi definida em duas classes, a saber: 1 – longa distância e 2 – outras.

Para a realização desta detecção, utilizou-se uma Rede Neural Artificial de múltiplas camadas MLP (*Multi Layer Perceptron*). Existem inúmeros atributos que podem ser utilizados na classificação da tomada, tais como o percentual de grama da imagem (recuperado por meio da segmentação descrita na Seção 4.1), além da quantidade de linhas existentes na imagem (informação recuperada por intermédio da Segmentação das Linhas do Campo descrita na Seção 4.2). A seguir, são listados todos os atributos utilizados na investigação.

1. Percentual de Grama da Imagem;
2. Número de Linhas;
3. Coeficiente Angular da Maior Reta;
4. Coeficiente Linear da Maior Reta;
5. Relação: Maior Linha/Área da Imagem;
6. Número de *Blobs* (Área homogênea de *pixels* de cor diferente do gramado);
7. Número de *Blobs* Filtrados;
8. Maior *Blob*;
9. Menor *Blob*;

O primeiro atributo escolhido, o percentual de grama da imagem, é relevante para a classificação de tomadas, visto que uma tomada será considerada de longa distância quando possuir grande quantidade de grama.

O número de linhas é outro atributo significativo, pois uma grande quantidade de linhas, em geral, é indício de que o quadro é de longa distância; um número pequeno de linhas, inclusive zero, faz com que o quadro tenha grande chance de pertencer à classe outras.

Os atributos 3, 4 e 5 são correlacionados, pois em uma imagem de longa distância as linhas serão muito mais evidentes, estarão em maior quantidade e, muito provavelmente, deverão ajudar na detecção da tomada de longa distância. Quando uma tomada não pertence a esta classe, tal como a filmagem da torcida e o *close-up* em jogadores, geralmente não apresentam linhas nos quadros de vídeo.

O atributo 6, Número de *Blobs*, também é relevante para a classificação de tomadas. Em tomadas de longa distância muitos jogadores devem estar presentes na imagem enquanto que em uma tomada rotulada como outras, este cenário em geral não ocorre. Devido a algumas falhas provenientes da Segmentação do Gramado, a característica Número de *Blobs* precisou ser adaptada para excluir *blobs* muito pequenos, que não são de interesse da aplicação, bem como a exclusão de *blobs* muito grandes,

e.g., a área exterior ao gramado. Por fim, as características 8 e 9 também são relevantes, haja vista que tomadas de longa distância possuem *blobs* relativamente pequenos, enquanto na classe outras esta característica não é obrigatória.

Uma vez definidas as características relevantes para a classificação de tomadas, os atributos mais relevantes foram encontrados por intermédio ao algoritmo *Best First* e uma Rede Neural Artificial foi treinada, para que esta classificação fosse feita de forma automática.

4.4. Detalhes de Implementação

Nas seções anteriores, foram apresentadas abordagens para a segmentação do gramado, segmentação das linhas do campo e classificação de tomadas. Visando à realização dos resultados experimentais descritos no Capítulo 5, foi necessário implementar a extração de todas as características utilizadas.

A implementação foi realizada utilizando-se a linguagem C/C++ a partir do ambiente de desenvolvimento integrado (*Integrated Development Environment* - IDE) Microsoft Visual Studio 2008. Dentre os fatores mais importantes para a escolha desta linguagem, está a existência de bibliotecas que auxiliam na manipulação de imagens, tais como *OpenCV* (BRADSKY; PISAREVSKY; BOUGUET, 2006) e desempenho superior em relação a velocidade de processamento quando comparadas a outras linguagens.

4.4.1. Segmentação das Linhas do Campo

A implementação da transformada de *Hough* foi auxiliada pelo uso da biblioteca de Visão Computacional *OpenCv*, por intermédio da função específica *cvHoughLines2(...)* que necessita dos 8 parâmetros descritos a seguir.

1. *CvArr* image;*
2. *void* line_storage;*
3. *int method;*
4. *double rho;*

5. double *theta*;
6. int *threshold*;
7. double *param1*;
8. double *param2*;

O primeiro parâmetro, *image*, corresponde a imagem na qual se deseja realizar a detecção das linhas, *line_storage* é um ponteiro que armazenará as informações de linhas recuperadas da imagem. O terceiro parâmetro é o *method* que pode ser: CV_HOUGH_STANDART, CV_HOUGH_PROBABILISTIC e CV_HOUGH_MULTI_SCALE.

Os parâmetros *rho* e *theta*, que definem a resolução do acumulador do plano, foram os mais complicados para calibrar, por serem bastante sensíveis. O *threshold* é responsável por definir o valor que limitará o acumulador de *pixels*. Por fim, os parâmetros *param1* e *param2* definem, respectivamente, o tamanho mínimo para que um segmento de reta seja recuperado pela transformada e a distância as linhas colineares.

A calibração dos parâmetros foi realizada de maneira subjetiva, por análise visual, por intermédio do desenvolvimento de uma interface que exibe, em tempo real, o resultado da aplicação da transformada de *Hough* e permite controlar o valor dos parâmetros específicos, conforme mostrado na Figura 2.15.

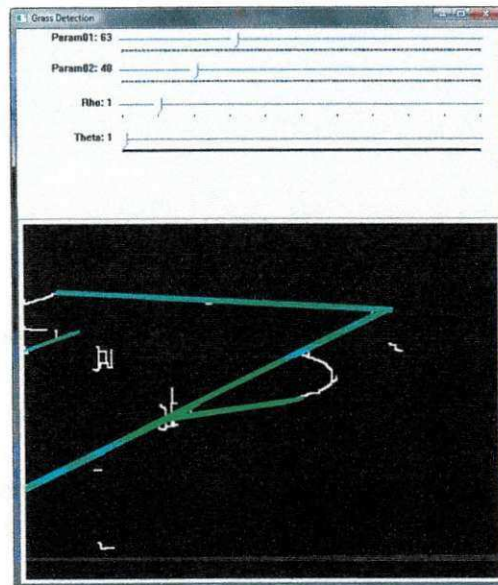


Figura 2.15 - Interface desenvolvida para auxiliar a escolha dos parâmetros da Transformada de *Hough*.

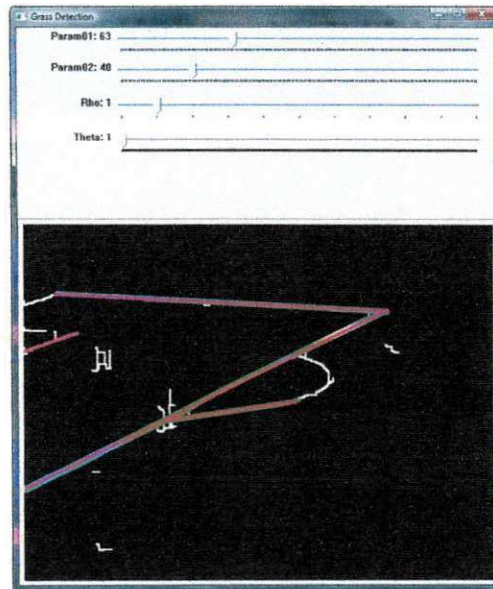


Figura 2.16 – Detecção de Linhas após o algoritmo de eliminação das linhas redundantes.

A aplicação do algoritmo de eliminação de linhas redundantes pode ser verificada na Figura 2.16, bem como o contraste entre a Transformada de *Hough* sem a eliminação das linhas (linhas verdes, Figura 2.15) e a transformada após a aplicação da redução (linhas vermelhas, Figura 2.16).

4.4.2. Classificação de Tomadas

Para que a Classificação de tomadas fosse possível, diversos atributos tiveram que ser extraídos da imagem para o treinamento da rede neural, alguns descritos em seções anteriores (Percentual de Grama e Número de Linhas) e outros, tais como Número de *Blobs* Filtrados, tiveram auxílio da biblioteca *OpenCV* (BRADSKY; PISAREVSKY; BOUGUET, 2006) para serem implementados.

A implementação da característica Número de *Blobs* foi desenvolvida por intermédio da função específica *cvFindContours(...)* que recupera os contornos existentes na imagem fornecida. A função utiliza os 6 parâmetros descritos a seguir.

1. `IplImage* img;`
2. `CvMemStorage* storage;`
3. `CvSeq** firstContour;`

4. *Int headerSize;*
5. *CvContournRetrievalModel mode;*
6. *CvChainApproxMethod method;*

O atributo *img* representa a imagem de entrada que deve ser binarizada. O atributo *storage* é responsável por armazenar as informações importantes dos contornos recuperados pelo método tais como localização e área do contorno. O parâmetro *FirstContourn* é um ponteiro para todos os contornos recuperados. O parâmetro *headerSize*, que tem como padrão *sizeof(CvContour)*, é o tamanho máximo para o contorno. Por fim, os parâmetros *mode* (o que fazer após terminar) e *method* (como fazer o cálculo da área) são definidos, respectivamente, como *CV_RETR_LIST* (recupera todos os contornos e salva em uma lista) e *CV_CHAIN_APPROX_SIMPLE* (comprime segmentos horizontais, verticais e diagonais, mantendo apenas seus pontos terminais).

Após a detecção de contornos, tornou-se imprescindível a filtragem de alguns *blobs* gerados por falhas na segmentação do gramado. A principal finalidade desta filtragem foi eliminar ruídos (*blobs* pequenos) e *blobs* muito grandes, restando apenas os que possivelmente fossem jogadores em tomadas de longa distância. Por intermédio da análise visual em diversos vídeos, foi possível verificar a relação entre a área de um jogador em tomada de longa distância e a área da imagem (ver Equação (4.7)) finalizando, assim, a obtenção de características.

$$0.37\% \leq \text{ÁreaDoJogador} / \text{ÁreaDa Imagem} \leq 3.02\% \quad (4.7)$$

Tanto para a seleção das características quanto para o treinamento da Rede Neural Artificial, foi utilizada a ferramenta *Weka*. O primeiro passo foi realizado por intermédio da funcionalidade *Select Attributes* → *Best First* (Figura 2.17). O algoritmo implementado no método *Best First* é uma busca que ordena todas as soluções parciais de acordo com algumas heurísticas e tenta prever como gerar uma solução completa. Desta forma, o método selecionou os melhores atributos para a Rede Neural, que foram: Percentual de Grama, Número de Linhas e Número de *Blobs* Filtrados.

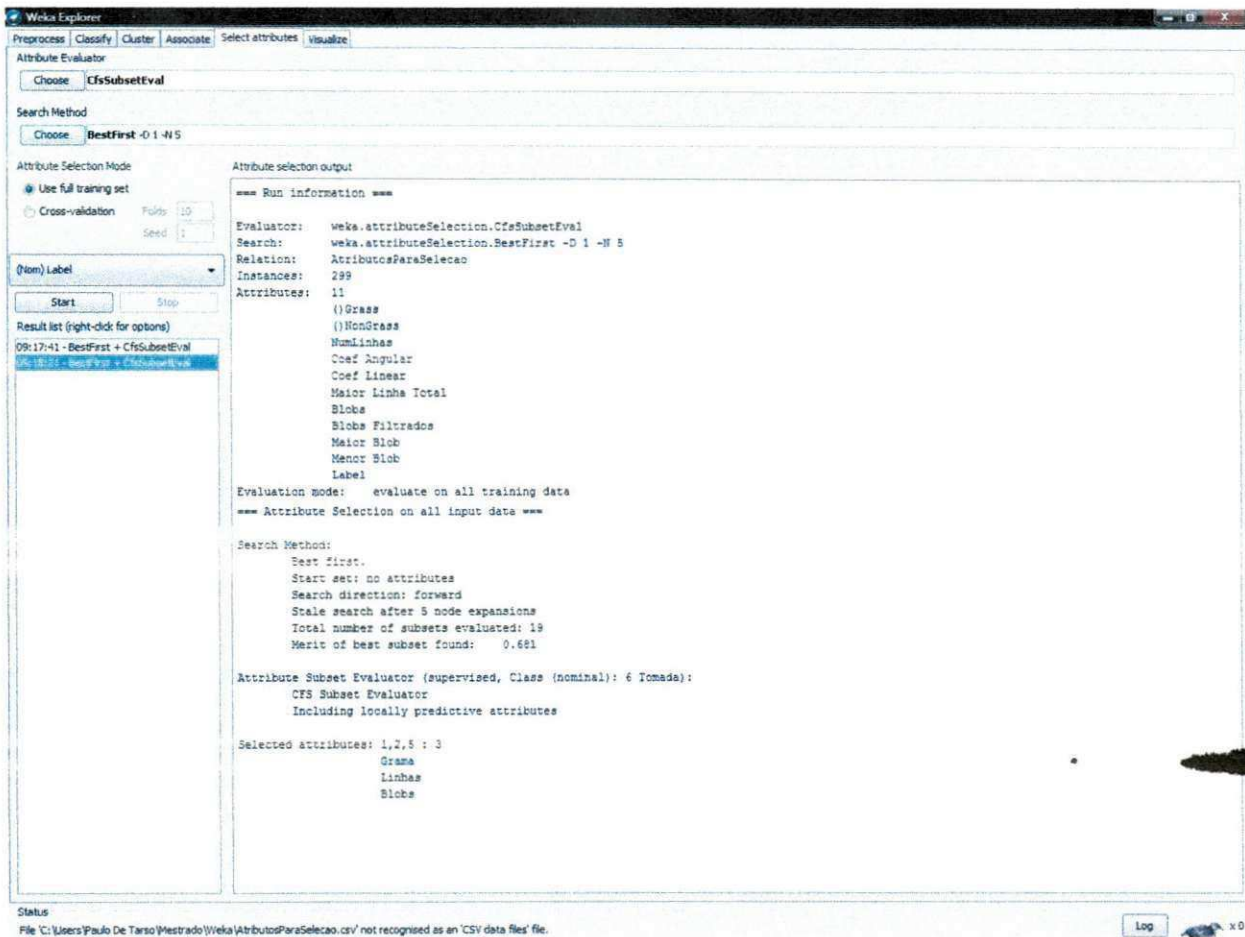


Figura 2.17 – Ferramenta Weka realizando a seleção de atributos mais importantes.

Para que a Rede Neural fosse a mais genérica possível, diversas combinações foram testadas no *Weka*, buscando-se as melhores taxas de classificação. Os melhores resultados foram obtidos por uma Rede MLP, funcionalidade *Classify* → *MultilayerPerceptron* (Figura 2.18), com semente igual a 30 (trinta), 2 camadas escondidas com 2 neurônios cada uma. A forma de treinamento utilizada foi a validação cruzada, em que o conjunto total foi dividido em 10 subconjuntos com o mesmo tamanho (*folds*). Para cada um dos 10 subconjuntos, a ferramenta realizou o treinamento com nove subconjuntos e o teste com o subconjunto restante.

4.5. Considerações Finais do Capítulo

Neste capítulo, foram propostas abordagens para a Segmentação do Gramado, a Segmentação das Linhas do Campo e a Classificação de

tomadas. Os principais passos realizados, as equações utilizadas e os detalhes de implementação, também foram apresentados no decorrer deste capítulo.

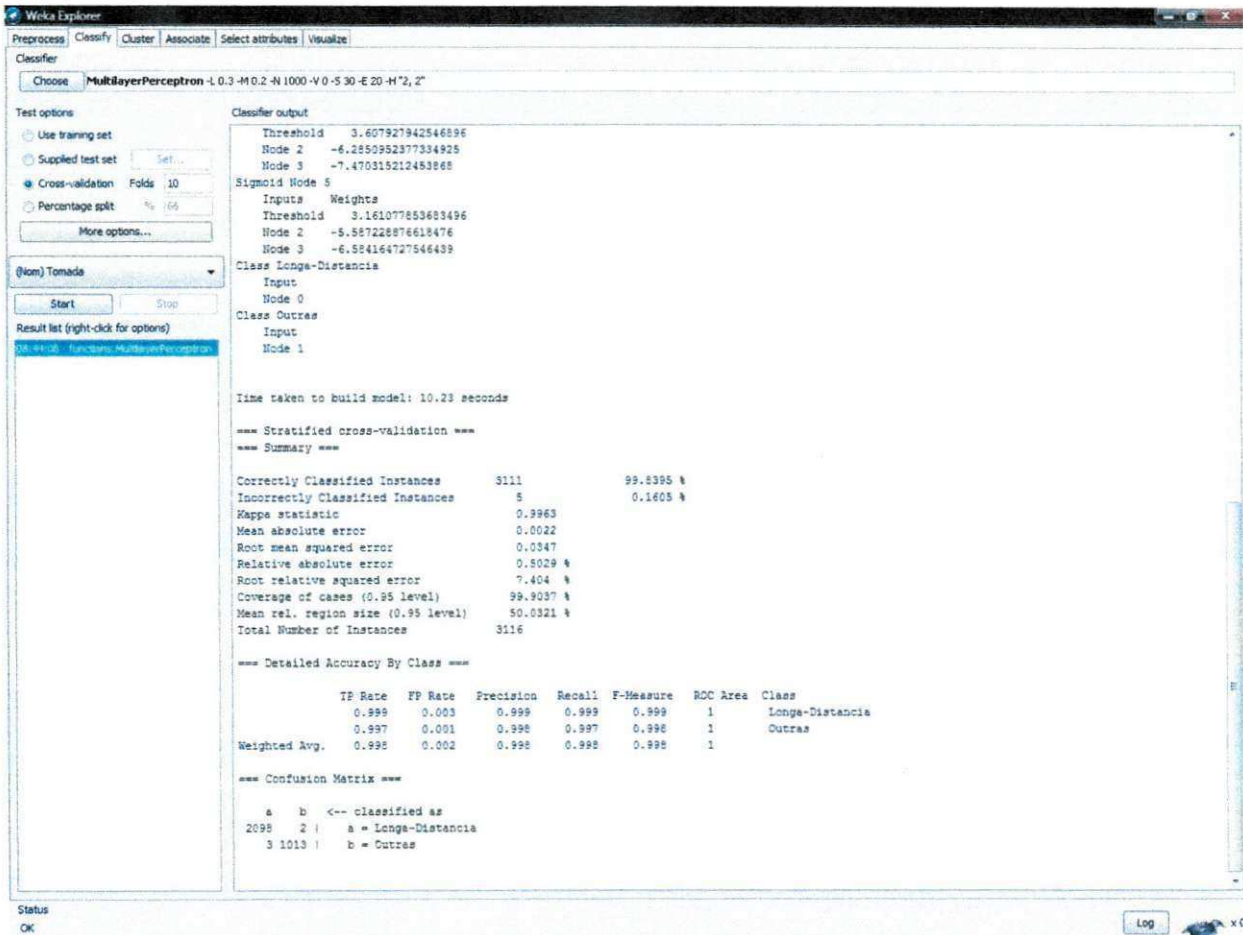


Figura 2.18 – Ferramenta Weka realizando o treinamento da rede neural.

A Segmentação do Gramado foi dividida nas seguintes sub-etapas: (i) obtenção da base; (ii) obtenção do mapa da imagem em nível de cinza, por meio do algoritmo de vizinhos mais próximos; e (iii) obtenção dos melhores limiares de binarização e número de imagens utilizadas na comparação.

Na abordagem para a Segmentação das Linhas, foram aplicadas máscaras visando à binarização da imagem e, em seguida, aplicada a Transformada de *Hough*; por fim, foi desenvolvido um algoritmo baseado em heurísticas, para eliminar as linhas redundantes provenientes desta transformada.

Concluindo as abordagens propostas, utilizou-se uma Rede Neural Artificial MLP (*Multi Layer Perceptron*) com os atributos extraídos na Segmentação do Gramado e das Linhas, para realizar a Classificação de tomadas, considerando as classes longa distância e outras.

No próximo capítulo, apresenta-se a avaliação experimental e os resultados obtidos com as abordagens desenvolvidas. No capítulo, descrevem-se todos os passos realizados na avaliação, apresentam-se gráficos de avaliação dos segmentadores treinados com auxílio da curva ROC e resultados obtidos pelo classificador de tomadas, por intermédio da ferramenta *Weka*.

Capítulo 5

Avaliação Experimental

5.1. Segmentação do Gramado

Objetivando uma investigação com respeito aos limiares utilizados por cada algoritmo, realizou-se uma análise quantitativa baseada no conceito de curva ROC¹ (PRATI; BATISTA; MONARD, 2008). As curvas ROC foram escolhidas devido à existência de limiares ou intervalos de limiares associados aos algoritmos, fato que, certamente, influencia o resultado final.

A curva ROC é utilizada para visualizar, avaliar, organizar e selecionar classificadores, levando-se em consideração seus desempenhos. Para realizar essas análises, gráficos ROC mostram a relação entre taxas de acerto (positivos verdadeiros) e taxas de alarmes falsos (falsos positivos) dos classificadores (PRATI; BATISTA; MONARD, 2008). Com referência ao problema abordado, ou seja, a segmentação binária da imagem em grama e não grama existe quatro saídas possíveis, que podem ser observadas na Tabela 5.1.

Para a avaliação do trabalho de Ekin, Tekalp e Mehrotra (2003), foi realizada a variação do limiar *t-color*, com fins à construção da curva ROC. Como o *t-color* é o principal responsável pela decisão do *pixel* ser ou não grama, tornou-se necessária a sua variação. Nesta dissertação, aquele limiar foi variado de 0 a 200, com saltos de 10, para que a curva ROC fosse obtida por completo. Para a avaliação do trabalho de Pallavi et al. (2008), que utiliza dois intervalos de valores no espaço cromático YIQ, também foi realizada a

¹ Sigla que significa: *Receiver Operating Characteristic*

variação desses intervalos. Como a faixa de variação de cores em I compreende o intervalo [-151,98; +151,98] e em Q o intervalo [-133,365; +133,365], é necessária a variação de todas as possíveis combinações, para que a curva ROC seja construída de forma mais contínua. Por fim, para a avaliação do trabalho de Cardoso (2009), foi variado o número de níveis de cinza da imagem, a fim de obter a curva ROC. Como apenas com o intervalo de variação atingindo o valor 16 a curva ROC já foi construída por completo, consideraram-se apenas os seguintes valores para os níveis de cinza: 0, 32, 64, 96, 128, 160, 192, 224 e 256.

Tabela 5.1 – Métricas utilizadas na análise ROC.

Métrica	Descrição
TP	Positivos Verdadeiros (True-Positive): são todos os <i>pixels</i> classificados pelo sistema como grama e que foram rotulados como tal no <i>ground-truth</i> .
FP	Positivos Falsos (False-Positive): são todos os <i>pixels</i> classificados pelo sistema como grama, mas que foram rotulados como não grama no <i>ground-truth</i> .
TN	Negativos Verdadeiros (True-Negative): são todos os <i>pixels</i> classificados como não grama e que foram rotulados como tal no <i>ground-truth</i> .
FN	Negativos Falsos (False-Negative): são todos os <i>pixels</i> que foram classificados como não grama, mas que foram rotulados como grama no <i>ground-truth</i> .
Taxa de TP	TPR = TP / (TP + FN): corresponde ao percentual de Positivos Verdadeiros com relação ao total de amostras positivas (que também correspondem à soma dos Positivos Verdadeiros com os Negativos Falsos).
Taxa de FP	FPR = FP / (FP + TN): corresponde ao percentual de Positivos Falsos com relação ao total de amostras negativas (que também correspondem à soma dos Positivos Falsos com os Negativos Verdadeiros).

Conforme comentado anteriormente, a curva ROC permite visualizar graficamente o efeito da variação de limiares de diferentes técnicas, por todo o seu espectro de valores, desde o valor mais restritivo até o valor mais liberal. É considerado ótimo aquele valor que faz a curva mais se aproximar das coordenadas (0%, 100%) (PRATI; BATISTA; MONARD, 2008). No método de Firmino Jr. e Gomes (2007), utilizado como base neste trabalho

de dissertação, dois são os elementos que podem interferir no desempenho do algoritmo: a quantidade k de vizinhos utilizados na comparação e o limiar I para a binarização após a imagem ser convertida para nível de cinza. Para obtenção dos resultados foi levado em consideração os 15 vídeos descritos na base de imagens, em que 10 foram utilizados para treinamento e 5 utilizados nos testes. Primeiramente, nesta dissertação, foi considerada apenas a quantidade k de vizinhos utilizados na comparação e então construída a curva ROC, apresentada na Figura 5.1.

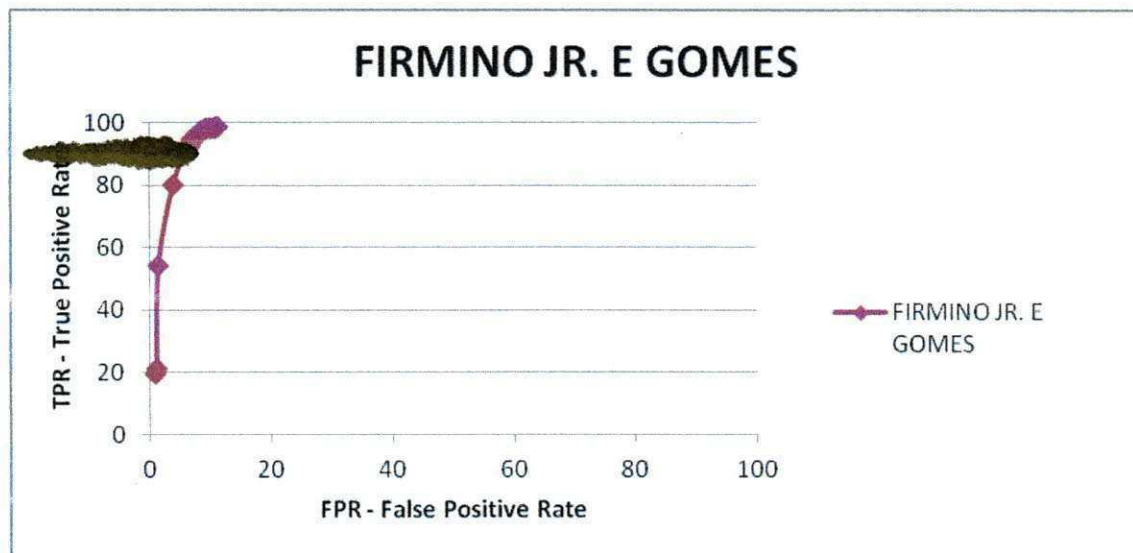


Figura 5.1 – Curva ROC para a variação dos k -vizinhos mais próximos.

De acordo com o gráfico da Figura 5.1, pode ser observado que, para o cenário dos k -vizinhos mais próximos, a quantidade de vizinhos igual a 5 (cinco) apresentou a menor distância para o ponto (0;100). Esta variação do número de vizinhos não preencheu completamente a área do gráfico, devido ao fato que uma quantidade muito grande de vizinhos utilizados no processamento poderia prejudicar de maneira considerável o desempenho do sistema. Fixando-se este valor e aplicando o mesmo procedimento para o limiar de binarização I , encontrou-se como valor ótimo o limiar 14 (quatorze). Finalizando o processo de comparação entre os algoritmos, foram implementados os três descritos no início da seção (EKIN; TEKALP;

MEHROTRA, 2003, PALLAVI ET AL. 2008 e CARDOSO 2009), realizada a variação dos seus parâmetros e construída a curva ROC que foram inseridas no mesmo gráfico, apresentado na Figura 5.2.

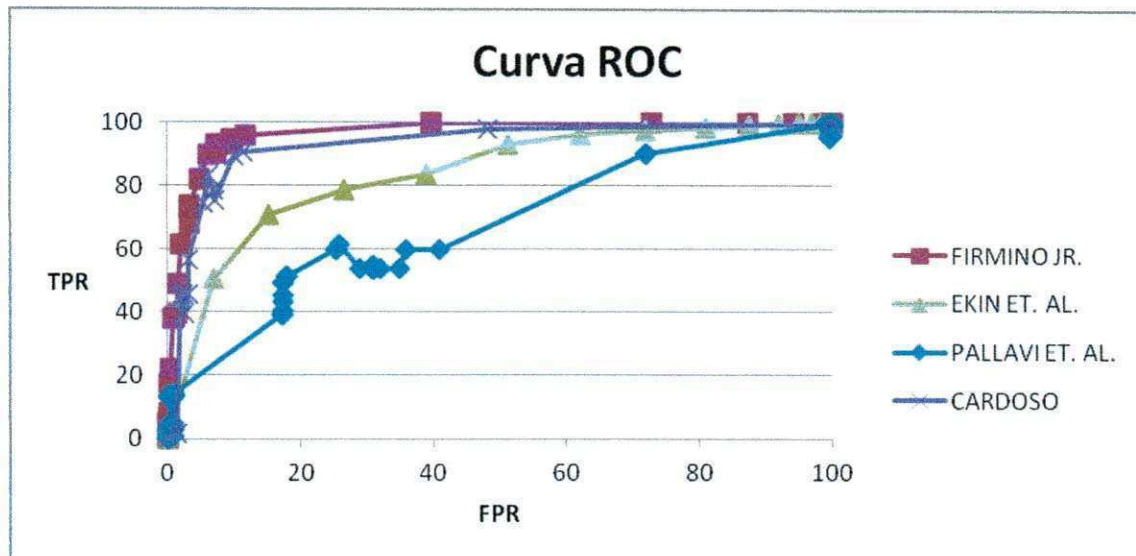


Figura 5.2 – Curva ROC dos quatro algoritmos abordados

Conforme se observa na Figura 5.2, a curva do algoritmo de Pallavi et al. (2008) apresenta comportamento irregular, devido ao fato de seu espaço de possibilidades não ter sido coberto totalmente. O algoritmo de Firmino Jr. e Gomes (2007) modificou apresentou o melhor resultado, pois segundo Prati, Batista e Monard (2008), quando comparadas às curvas ROC, caso não haja interseções, será considerada melhor aquela que se aproxima mais do ponto (0%; 100%). Um resultado do algoritmo modificado com relação à forma sistemática de obtenção dos parâmetros e aumento da base de imagens utilizada para a segmentação grama/não grama é mostrado na Figura 5.4. Adicionalmente, são exibidos também os resultados dos algoritmos comparados na curva ROC da Figura 5.1, em que se observa que o melhor resultado para o método proposto apresentou 98,32% de Taxa de Positivos Verdadeiros (TPR) e 16,08% de Taxa de Positivos Falsos (FPR).

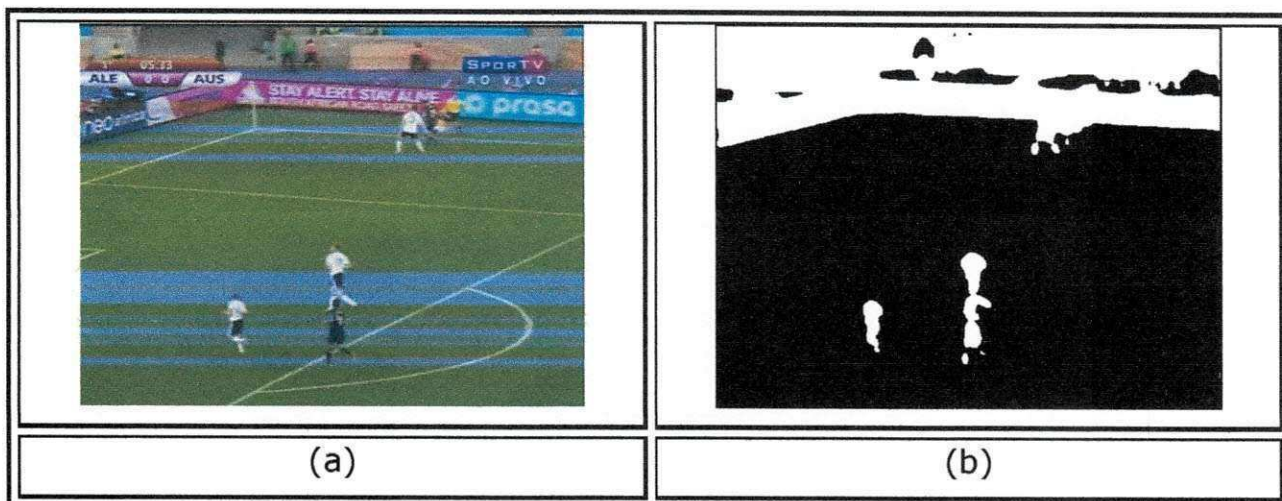


Figura 5.3 – Segmentação por intermédio do algoritmo Firmino Jr. E Gomes (2007) modificado: (a) Imagem original; (b) Imagem resultante.

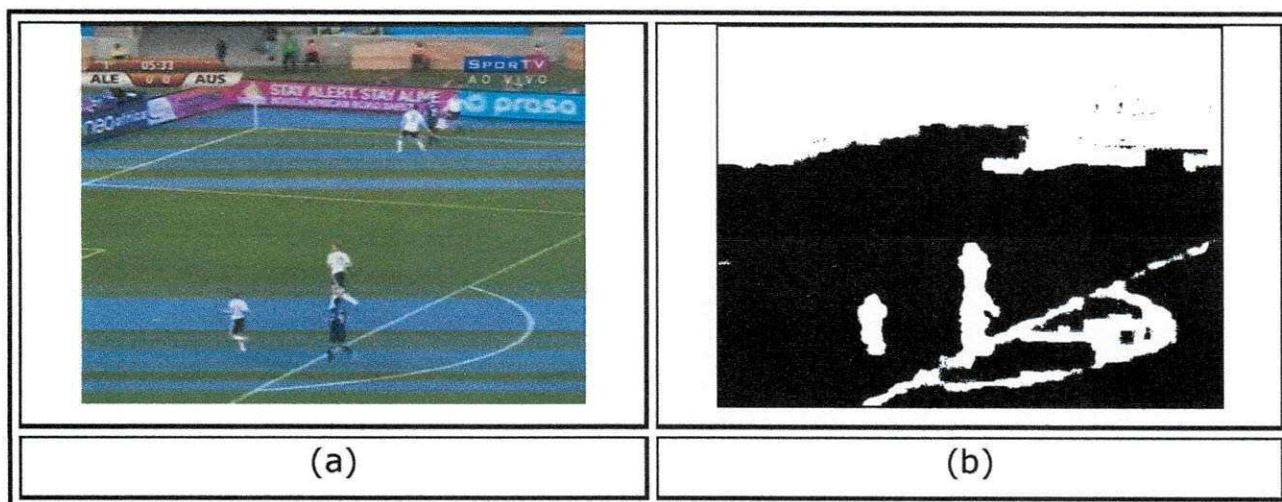


Figura 5.4 – Segmentação por intermédio do algoritmo Cardoso (2009): (a) Imagem original; (b) Imagem resultante.

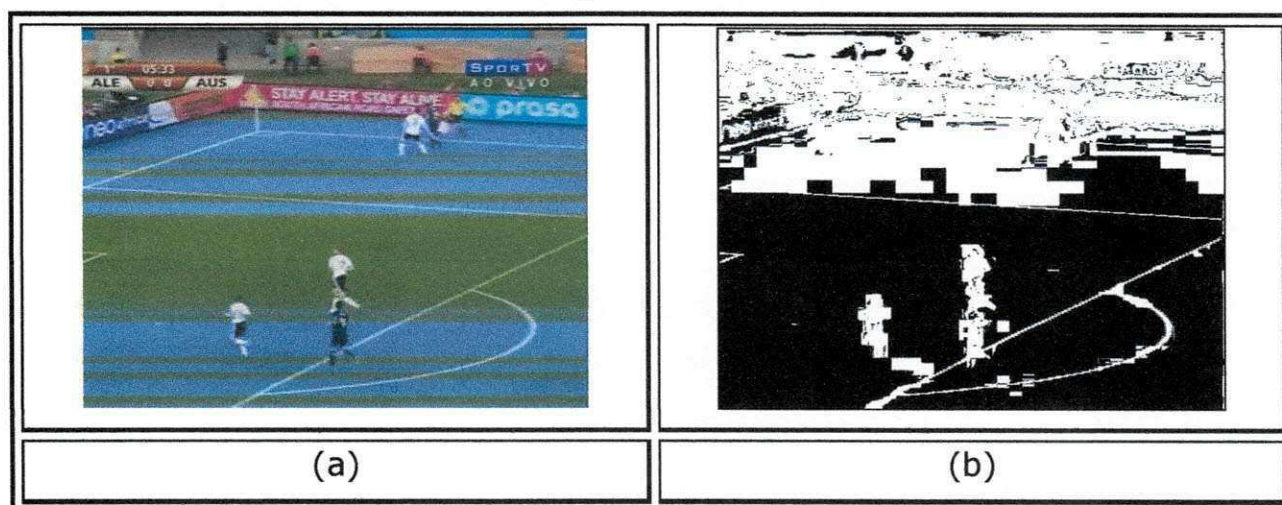


Figura 5.5 – Segmentação por intermédio do algoritmo Ekin, Tekalp E Mehrotra (2003): (a) Imagem original; (b) Imagem resultante.

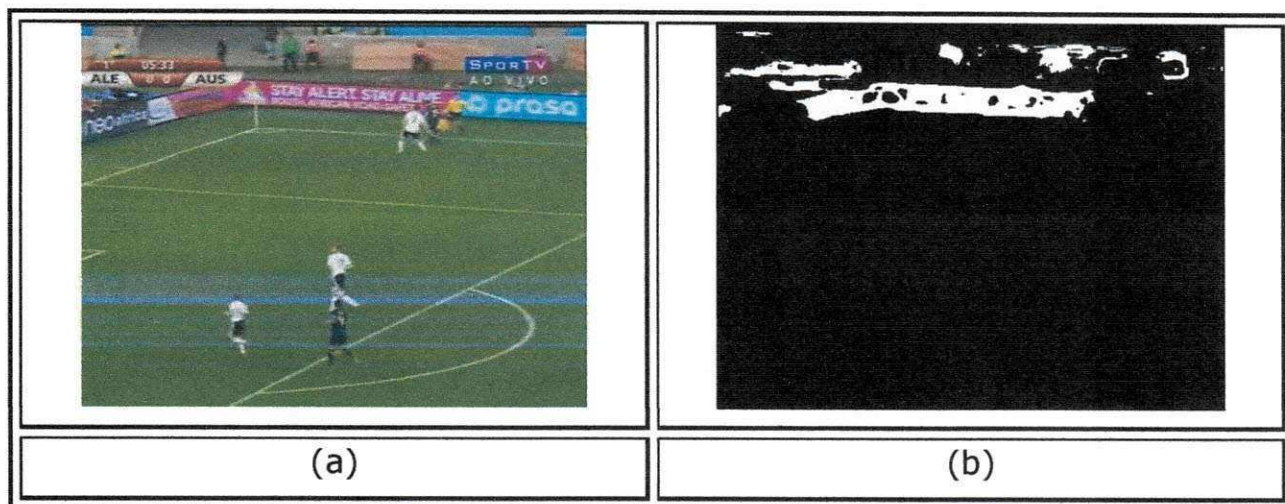


Figura 5.6 – Segmentação por intermédio do algoritmo Pallavi et al. (2008): (a) Imagem original; (b) Imagem resultante.

Conforme pode ser observado na Figura 5.3, o trabalho desenvolvido nesta dissertação ainda apresenta algumas falhas com relação à preservação das linhas do campo. Outra falha que o algoritmo desenvolvido ainda apresenta é a confusão entre o verde do gramado e o do uniforme de algumas equipes, conforme pode ser verificada no Jogo 03 (Apêndice A). Entretanto, nos Jogos 10 e 13, essa falha já não se apresenta tão evidente. Outro problema decorre do fato de presença de neblina em partidas, o qual compromete de forma significativa a segmentação do gramado, conforme foi constatado no Jogo 11. De maneira geral, em condições climáticas consideradas normais, o resultado visual apresentado no Apêndice A pode ser considerado satisfatório.

Por outro lado, a abordagem proposta por Cardoso 2009, (ver Figura 5.4), apresenta falhas na delimitação da região do gramado. Em Ekin et al. (2003) as linhas são corretamente recuperadas, mas falhas na região do gramado comprometem o resultado final. Por fim, em Pallavi et al. (2008) não se realiza corretamente a segmentação de jogadores, aumentando de forma significativa a taxa de Positivos Falsos.

5.2. Segmentação de Linhas do Campo

A maior parte dos trabalhos apresentados na revisão bibliográfica para a segmentação das linhas não apresenta uma avaliação experimental objetiva dos resultados. Um exemplo é o trabalho de Kamarposhty, Asadollahi e Asadollahi (2009) que exhibe apenas uma comparação visual com a transformada de *Hough* para algumas imagens. Em Sun e Liu (2009), no entanto, é realizado um experimento para verificar o tipo da linha detectada, dentre as classes Linha de meio-campo, linha de fundo e linha do pênalti. Em outros trabalhos, e.g., Cai e Tai (2005), são exibidos apenas resultados visuais. Por este motivo, aliados a não apresentação de detalhes suficientes para reprodutibilidade dos experimentos e a não existência de implementações disponíveis, não foi realizada uma análise comparativa entre a técnica proposta e as demais existentes na literatura.

A avaliação experimental desenvolvida nesta etapa da presente dissertação consiste em considerar a relação TPR (*True Positive Rate*) versus FPR (*False Positive Rate*), por meio da construção da curva ROC, da mesma forma que foi empregada na seção anterior. Para tanto, a imagem de referência (*ground-truth*), obtida da base de imagens e descrita na Seção 3.1 (Figura 5.7 (b)) é comparada com a imagem gerada após o processamento (Figura 5.7 (c)).

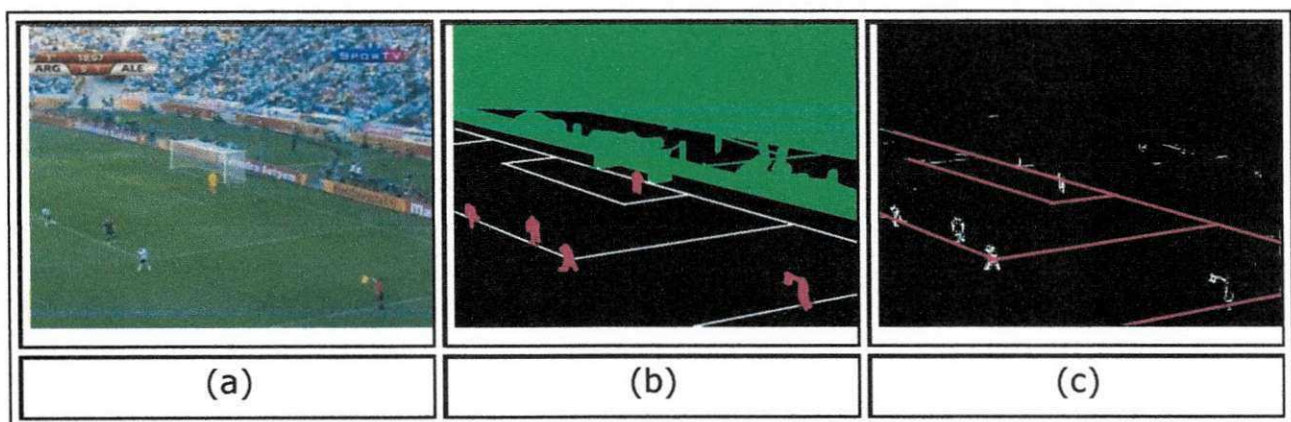


Figura 5.7 – Segmentação das Linhas: (a) Imagem original; (b) Imagem rotulada; (c) Imagem segmentada.

Ao se realizar esta avaliação experimental, constatou-se que a marcação da base considerou linhas mais espessas do que aquelas geradas pela segmentação, gerando taxas de erro que não retratam a realidade da segmentação desenvolvida. Na Figura 5.8, a marcação em cor verde representa os Falsos Negativos, enquanto a marcação em vermelho retrata os Falsos Positivos gerados pela implementação desenvolvida. Sendo assim, foi desenvolvida uma abordagem complementar para considerar a vizinhança do *pixel* selecionado, ou seja, caso o *pixel* segmentado na imagem processada esteja em uma vizinhança de tamanho X e este tenha sido marcado no *Ground-Truth*, este *pixel* será computado como correto em sua marcação. Neste trabalho, uma vizinhança de tamanho 5 mostrou bons resultados.

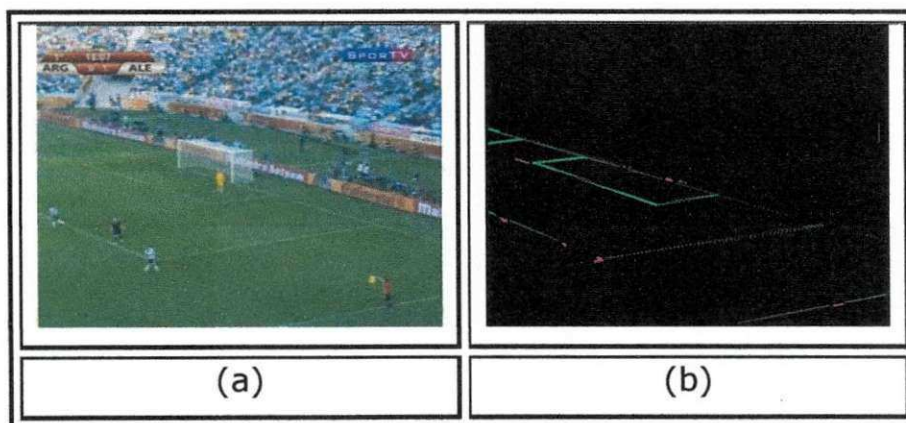


Figura 5.8 - Mapa de Erros da Figura 5.7.

Após a implementação desta abordagem complementar foi necessária a definição do melhor fator multiplicativo (Equação 5.1) para obtenção do limiar de binarização. Por intermédio da Curva ROC, os limiares foram variados sendo considerado melhor aquele que se aproxima mais do ponto (0%; 100%) Prati, Batista e Monard (2008) que, no gráfico da Figura 5.9, foi o ponto associado ao valor 2,3. A fórmula para a obtenção do limiar pode ser verificada na Equação 5.1.

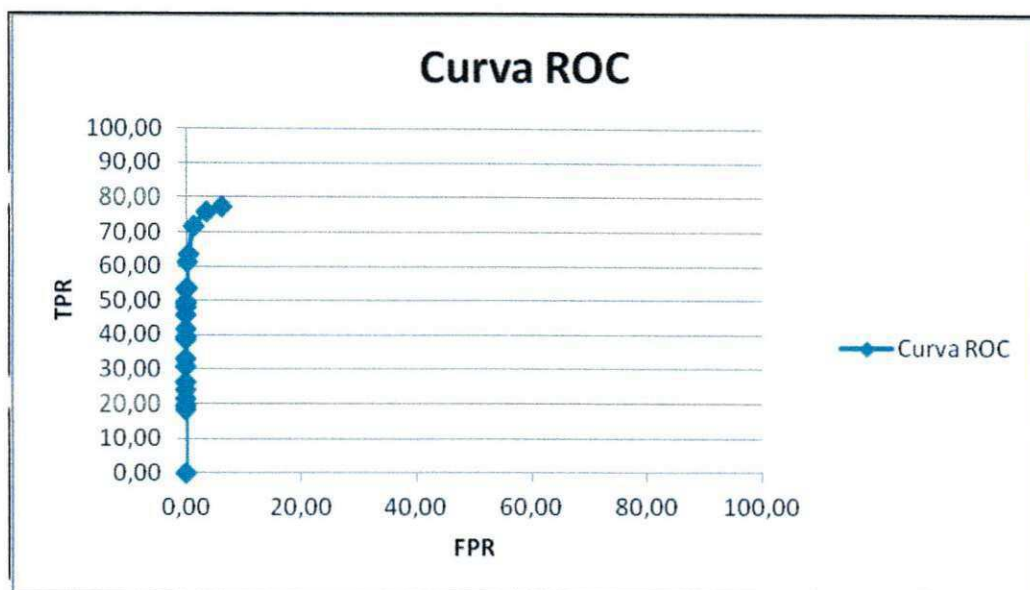


Figura 5.9 – Curva ROC para a Segmentação das Linhas.

$$LB = IntensidadeDoPico * 2.3 \quad (5.1)$$

Apesar de se ter realizado a variação do fator multiplicativo para a obtenção do limiar de binarização, não foi possível considerar todos os valores na construção da Curva ROC, fato que só seria possível caso os parâmetros da transformada de *Hough* fossem variados, pois mesmo que a imagem esteja segmentada em uma só cor (preta ou branca), o valor 100% não será atingido em ambas as taxas (*TPR* e *FPR*). Os resultados da segmentação das linhas para o melhor limiar podem ser verificados na Tabela 5.2.

Tabela 5.2 – Melhores Resultados da Segmentação de Linhas.

TPR (%)	FPR (%)
77,34	6,23

Alguns resultados visuais foram dispostos no Apêndice B, em que se pode constatar que a segmentação das linhas foi realizada com sucesso em linhas retilíneas (linhas 01, 03 e 11 da Tabela B.1), por conta da característica da transformada utilizada (*Hough Transform*). As falhas

existentes em linhas retas devem-se ao fato que em algumas imagens utilizadas as linhas são pouco visíveis (linha 14 da Tabela B.1) e, em outros casos, a neblina atrapalhou a correta segmentação (linha 13 da Tabela B.1).

5.3. Classificação de Tomadas

No trabalho desenvolvido por Yang et al. (2007), além de se detectar a posição do campo em que o lance está acontecendo, é realizada também a classificação da tomada em longa, curta ou média distância. A base utilizada nos testes é composta por aproximadamente 30 mil imagens de longa distância que foram recuperadas de 5 vídeos diferentes e as taxas de classificação (TPR) obtidas naquele trabalho estão em torno de 98,56%.

Outra abordagem para classificação de tomadas é apresentada em Ye et al. (2005), utilizando 15 vídeos e com TPR de 92,5%. Por fim, o trabalho apresentado por Lopes, Fukuma e Rodrigues (2011), o qual realiza a detecção de tomadas por meio da Teoria da Informação Não-Extensiva, apresenta precisão de até 97%, de acordo com a curva ROC.

O principal objetivo da classificação de tomadas desenvolvida neste trabalho foi detectar as tomadas consideradas de longa distância, que se caracterizam por grande quantidade de jogadores, grande quantidade de grama e jogadores em tamanho muito pequeno, em relação ao tamanho da imagem processada. Assim, a rotulagem foi feita considerando-se apenas duas classes: longa distância e outras, tal como em Lopes, Fukuma e Rodrigues (2011).

A avaliação experimental realizada neste trabalho foi composta, no total, de 8365 imagens de 4 diferentes vídeos, detalhadas na Seção 3.2. Com o auxílio da ferramenta *Weka* foram obtidas taxas superiores aos trabalhos existentes na literatura. Entretanto, estas são conclusões parciais, uma vez que a base de teste utilizada foi diferente daquelas utilizadas pelos outros autores e apresenta número menor de imagens, em alguns casos, pois a

base utilizada por Yang et al. (2007) possui cerca de 30 mil imagens de longa distância e a base de Lopes, Fukuma e Rodrigues (2011) cerca de 45 mil. Os resultados são mostrados na

Tabela 5.3.

Tabela 5.3 – Resultados da Classificação das Tomadas.

Classe	TPR (%)	FPR (%)	Área ROC (%)
longa distância	99,6	0,16	99,6
outras	98,4	0,04	99,6

Mesmo com uma alta taxa de acerto, algumas imagens não foram classificadas de forma correta. No Apêndice C se pode verificar que a principal causa para os Falsos Positivos, i. e., imagens da classe longa distância classificada de forma incorreta, foi o insucesso na segmentação dos *blobs*. Por sua vez, a causa para os Falsos Negativos, i. e., imagens da classe outras classificada de forma incorreta, foi o grande percentual de grama, fator decisivo para a incorreta classificação.

5.4. Considerações Finais do Capítulo

Neste capítulo, foi descrita a parte experimental do trabalho desenvolvido nesta dissertação. Esta parte experimental tem, como principal objetivo, validar as abordagens propostas e descritas no Capítulo 4. Após todos os experimentos realizados, pôde-se verificar que:

- A curva ROC do Segmentador do Gramado proposto apresentou resultados superiores àqueles apresentados pelos demais trabalhos existentes na área. Como na maioria dos trabalhos desenvolvidos a segmentação do gramado é um passo intermediário, apenas resultados visuais eram apresentados, lacuna que foi sanada neste

trabalho com a obtenção de resultados experimentais objetivos.

- No tocante à Segmentação das Linhas do Campo, os trabalhos revisados não apresentaram uma avaliação estatística ou então não forneciam informações suficientes para sua reprodução. Os resultados apresentados pela abordagem proposta foram considerados satisfatórios, como se pôde verificar a partir da curva ROC.
- Por fim, o Classificador de Tomadas apresentou resultados excelentes, chegando a atingir 100% de correta detecção em alguns vídeos e finalizando com o percentual de 99,14%, em toda a base de imagens utilizada.

Capítulo 6

Considerações Finais

Neste capítulo, são apresentadas: (i) as contribuições resultantes da pesquisa desenvolvida e (ii) a formulação de propostas para trabalhos futuros, levando-se em consideração as dificuldades e os problemas encontrados durante o desenrolar desta pesquisa.

6.1. Contribuições do Trabalho Realizado

A partir dos resultados que foram apresentados no Capítulo 5, foi possível concluir que o principal objetivo do trabalho, avaliar, implementar e testar técnicas existentes e propor inovações para Segmentação do Gramado e das Linhas do Campo e Classificação de Tomadas, foi atingido. Como forma de validação das técnicas desenvolvidas, uma base de imagens foi concebida e disponibilizada para que os novos algoritmos desenvolvidos sejam testados. Desta forma, as principais contribuições da pesquisa são:

1. Construção de uma Base de Imagens que abrange os três aspectos considerados no desenvolvimento deste trabalho;
2. Realização de análise objetiva de resultados em cada um dos aspectos, utilizando-se da curva ROC, e taxas TPR (*True Positive Rate*) e FPR (*False Positive Rate*);
3. Desenvolvimento de uma extensão de uma abordagem existente para Segmentação de Grama, considerando-se aspectos de vizinhos mais próximos no modelo de cores RGB;

4. Avaliação dos melhores parâmetros utilizados na Transformada de *Hough*, para detecção das linhas do campo;
5. Construção de um sistema para gerenciar as principais informações relacionadas às linhas detectadas na imagem;
6. Obtenção de um Classificador de Tomadas (longa distância vs outras) que apresenta resultados experimentais relativamente melhores que aqueles documentados na literatura. Entretanto, este resultado foi obtido a partir de uma base de dados diferente daquela utilizada pelos demais autores.

Os principais resultados obtidos podem ser resumidos como:

- A curva ROC do Segmentador do Gramado proposto apresentou resultados superiores àqueles obtidos pelos demais segmentadores existentes analisados, comprovando seu desempenho superior.
- Os resultados apresentados pela abordagem proposta para a Segmentação das Linhas do Campo foram considerados satisfatórios, conforme verificado a partir da curva ROC.
- O Classificador de Tomadas apresentou resultados excelentes, chegando a atingir 100% de correta detecção em alguns vídeos e finalizando com o percentual de 99,14%, em toda a base de imagens utilizada.

Em seguida, são apresentadas sugestões para a melhoria e a complementação dos resultados obtidos nesta pesquisa.

6.2. Trabalhos Futuros

Durante o desenvolvimento do presente trabalho, várias lacunas de pesquisa ainda não foram preenchidas, principalmente por conta da complexidade no desenvolvimento de algumas técnicas e pelo fato da existência de muitos elementos importantes em uma partida de futebol.

As taxas obtidas pelo método proposto para a Segmentação das Linhas do campo ficaram abaixo das taxas obtidas com os métodos propostos para os demais problemas investigados nesta dissertação. Como existem várias linhas não retilíneas num campo de futebol, provavelmente a aplicação da Transformada de *Hough* Circular proposta em Atiquzzaman (1992) poderá trazer melhores resultados.

Outro passo importante que pode ser realizado a partir dos segmentadores e do classificador desenvolvido é a análise de vídeos e não apenas imagens estáticas, como aquela que foi realizada neste trabalho. Assim, informações temporais poderão ser adicionadas aos segmentadores/classificadores e os resultados obtidos poderão ser melhorados.

A detecção da bola (Seção 2.2) e a detecção de jogadores (Seção 2.3) são de fundamental importância para a boa análise de uma partida de futebol. O primeiro passo, a detecção de *blobs*, já foi realizado no presente trabalho (Seção 3.2) e pode ser melhorado/modificado para realizar a detecção de jogadores.

Referências Bibliográficas

ARAÚJO, E. R. de - Revocação (recall) e precisão (precision) no SDI/CIN/CNEN. *Ciência da Informação*, v. 8, n. 1, p. 47-50, 1979.

ARNAUT, R. D.; COSTA, A. H. R. Rastreamento de jogadores de futebol em seqüência de imagens. *Revista Eletrônica TECCEN*, v. 2, n. 1, p. 44-53, 2009.

ASSFALG, J.; BERTINI, M.; COLOMBO, C.; BIMBO, D. A.; NUNZIANI, W. Semantic annotation of soccer videos: automatic highlights identification. *Computer Vision and Image Understanding, Special issue on video retrieval and summarization*, Estados Unidos, v. 92, n. 2, p. 285-305, 2003.

ATIQUZZAMAN, M. Multiresolution Hough transform - an efficient method of detecting patterns in images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 14, n. 11, p. 1090-1095, 1992.

BEETZ M.; GEDIKLI S.; BANDOUNCH J.; KIRCHLECHNER B., HOYNINGEN-HUENE N. V.; PERZYLO A. Visually Tracking Football Games Based on TV Broadcasts. *Int. Joint Conference on Artificial Intelligence IJCAI*, n. 20, v.1, p. 2066-2071, 2007.

BRADSKY, G. R.; PISAREVSKY, V.; BOUGUET, J. Learning OpenCV: Computer Vision with the OpenCV Library. First Edition. Springer: O ´REILLY, 2006.

BRINKHOFF, T.; KRIEGEL, H. P.; SCHNEIDER, R. Comparison of approximations of complex objects used for approximation-based query processing in spatial database systems. *In the Proceedings of 9th International Conference on Data Engineering*, p. 40-49, 1993.

CAI, Z. Q.; TAI, J. Line detection in soccer video. *IEEE International Conference on Information, Communications and Signal Processing (ICICS)*, v. 1, n. 5, p. 538 – 541, 2005.

CARDOSO, F. H. B. *Um método Automático para o Ajuste de Segmentadores de Imagens Baseados em Informação de Textura e Cor*. 2009. 126f. Dissertação (Mestrado em Computação) - Universidade Federal de Campina Grande, Campina Grande.

CHESTER, E. *Asus O!Play HDP-R1 HD Media Player*. Disponível em: <http://www.trustedreviews.com/Asus-O-Play-HDP-R1-HD-Media-Player_Peripheral_review>. Acesso em 17/01/2012.

DE BEM, R. A. *Uma abordagem livre de modelo para rastreamento de objetos em seqüência de imagens*. 2007. 84f. Dissertação (Mestrado em Engenharia Elétrica) - Escola Politécnica, Universidade de São Paulo, São Paulo.

D'ORAZIO, T.; GUARAGNELLA, C.; LEO, M.; DISTANTE, A. A new algorithm for ball recognition using circle Hough transform and neural classifier. *Pattern Recognition*. v. 37, n. 1, p. 393-408, 2004.

DUARTE, G. D. Uso da Transformada de Hough na Detecção de Círculos em Imagens Digitais. *Thema Revista Científica do Centro Federal de Educação Tecnológica*, Pelotas, v. 4, n. 1, p. 51-58, 2003.

EKIN, A.; TEKALP A. M.; MEHROTRA, R. A framework for tracking and analysis of soccer video. *Symposium Electronics Imaging: Science and Technology: Visual Communication and Image Processing*. v. 10, n. 5, p. 763-774, 2002.

EKIN, A.; TEKALP A. M.; MEHROTRA, R. Automatic Soccer Video Analysis and Summarization. *IEEE Transactions on Image Processing*. v. 12, n. 7, p. 796-807, 2003.

FIGUEROA, P. J.; LEITE, N. J.; BARROS, R. M. L. Tracking soccer players aiming their kinematical motion analysis. *Computer Vision and Image Understanding*, v. 10, n. 2, p. 122-135, 2006.

FIRMINO JR., P. T.; GOMES, H. M. Detecção de lances no campo de ataque em uma partida de futebol. *In Anais do Simpósio Brasileiro de Computação Gráfica e Processamento de Imagens 2007*, p.34-38, 2007.

FISHER, R.; PERKINS, S.; WALKLER A.; WOLFART E. Line Detection. Disponível em: <<http://homepages.inf.ed.ac.uk/rbf/HIPR2/linedet.htm>>. Acesso em 02/02/2012.

GIESSEL, R. *Joinville no YouTube*. Disponível em: <<http://gazetadejoinville.blogspot.com/2009/03/joinville-no-youtube.html>>. Acesso em 26/11/2010.

GONZALEZ, R. C.; WOODS, R. E.; Digital Image Processing. Third Edition. Estados Unidos: Pearson Prentice Hall, 2008. 954 p.

KAMARPOSHTY, M.S.; ASADOLLAHI, H.; TEYMOORI, M.M. Line Detection in Sport Images Using Multi-agent Systems. *Proceedings of the International Conference on Future Computer and Communication*, Kuala Lumpur, v. 2, p. 319 – 323, 2009.

KHATOONABADI, S. H.; RAHMATI M. Automatic soccer players tracking in goal scenes by camera motion elimination. *Image and Vision Computing*, New Zealand, v. 27, n. 4, p. 469-479, 2009.

LEVINE, D. M.; BERENSON, M. L.; STEPHAN, D. Estatística: Teoria e aplicações. Rio de Janeiro: LTC Editora, p. 374, 2004.

LOPES, G.; FUKUMA, W.; RODRIGUES, P. Detecção de Tipos de Tomadas em Vídeos de Futebol Utilizando Teoria da Informação Não-Extensiva. *VII Workshop de Visão Computacional - Sessão de Pôsteres*, Curitiba, Paraná, p. 443-448, 2011.

LOTFI, E.; POURREZA, H. R.; Summarizing Soccer Videos without Detecting the Events. *IJCSNS International Journal of Computer Science and Network Security*, v.6, n.10, 2006.

MILLS, S.; PRIDMORE, T.; HILLS, M. Tracking in a hough space with the extended kalman filter. *The British Machine Vision Conference*, Norwich, p. 173-182, 2003.

PALLAVI, V.; MUKHERJEE, J.; MAJUMDAR, K.; SURAL, S. Ball detection from broadcast soccer videos using static and dynamic features. *Journal of Visual Communication and Image Representation*, v. 19 n. 7, p. 426-436, 2008.

PEI, C.; YANG, S. Y.; GAO, L.; MA, W. P. A Real Time Ball Detection Framework For Soccer Video. *WSSIP0*, v. 15, p. 1-4, 2009.

PEPERAIO, L. *O fim do Google Vídeo*. Disponível em: <<http://www.lucaspeperaio.com.br/blog/o-fim-do-google-video>>. Acesso em 02/02/2012.

PRATI, R. C.; BATISTA, G. E. A. P. A.; MONARD, M. C. Curvas ROC para a avaliação de classificadores. *Revista IEEE América Latina*, São Paulo, 2008. Disponível em: <<http://www.icmc.usp.br/~gbatista/publications.html>>. Acesso em: 20 dez. 2008.

PRESS, G. *Quase metade do planeta sintonizado em casa na África do Sul 2010*. Disponível em: <<http://pt.fifa.com/worldcup/archive/southafrica2010/organisation/media/newsid=1473439>>. Acesso em 02/02/2012.

SALES, P.; *Viterbi Algorithm for Decoding of Convolutional Codes*. Disponível em: < <http://www.1-core.com/library/comm/viterbi/viterbi.pdf> >. Acesso em 22/03/2012.

SANTOS, M. *Tv Digital: O surgimento*. Disponível em: <<http://focamar.wordpress.com/2007/12/04/tv-digital-o-surgimento/>>. Aceso em 21/02/2011.

SEO, K.; KO, J.; AHN, I.; KIM, C. An Intelligent Display Scheme of Soccer Video on Mobile Devices. *IEEE Transactions on Circuits and Systems for Video Technology*, v. 17, n. 10, p. 1395-1401, 2007.

SGARBI E., BORGES L. D. Structure in Soccer Videos: detecting and classifying highlights for automatic summarization. *Iberoamerican Congress on Pattern Recognition - CIARP*, Cuba, v. 10, n. 5, p. 691-700, 2005.

SILVA, R. *Sky Digital: 9,3 milhões de assinantes*. Disponível em: <<http://www.telesatelite.net/novidades/tp.asp/tp/9759/sky-digital-9-3-milhoes-de-assinantes.html>>. Acesso em 17/01/2012.

SPAGNOLO, P.; MOSCA, N.; NITTI, M.; DISTANTE, A. An Unsupervised Approach for Segmentation and Clustering of Soccer Players. *International*

Machine Vision and Image Processing Conference, Irlanda, v. 08, n. 7, p. 133-142, 2007.

SUN, L.; LIU G. Field lines and players detection and recognition in soccer video. *ICASSP*, Taiwan, v. 1, n. 9, p. 1237-1240, 2009.

TRUONG, B. T.; VENKATESH, S. Video abstraction: A systematic review and classification. *ACM Trans. Multimedia Comput. Commun. Appl.*, v. 3, p. 3, 2007.

VIVIANI, O. *Fifa reabre teste de uso da tecnologia*. Disponível em: <http://www.jornalpequeno.com.br/2010/10/21/fifa-reabre-teste-de-uso-da-tecnologia-135489.htm>. Acesso em: 30/01/2011

WANG, J.; XU, C.; CNHG, E.; YU, X.; TIAN, Q. Event detection based on non-broadcast sports video. *Proc. of IEEE ICIP*, v. 01, p. 1-5, 2004.

YANG, B.; SUN, L.-F.; WANG, F.; WANG, P.; YANG, S.-Q. Mid-Level Descriptors Extraction of Soccer Video with Domain. *IEEE International Conference on Systems, Man and Cybernetics*, Taipei, v. 06, p. 4937 – 4941, 2007.

YE, Q.; HUANG, Q.; GAO, W.; JIANG, S. Exciting Event Detection in Broadcast Soccer Video with Mid-level Description and Incremental Learning. *ACM Multimedia*, Hilton, Singapore, v. 13, p. 455-458, 2005.

YOON, H.-S; BAE, Y.L.J.; YANG, Y.K. A soccer image sequence mosaicing and analysis method using line and advertisement board detection. *Electronics and Telecommunications Research Institute - ETRI*. v. 24, n. 6, 443-454, 2002.

YU, J. Q.; TANG, Y.; WANG, Z. F.; SHI, L. J. Playfield and ball detection in soccer video. *International Conference on Advances in Visual Computing*, Lake Tahoe, NV, USA, v. 3, n. 4, p. 387-396, 2007.

Apêndice A

Resultados da Segmentação do Gramado

Neste apêndice, apresentam-se exemplos de imagens após a aplicação da segmentação do gramado descrita na Seção 4.1. A Tabela A.1 apresenta o resultado final de 90 imagens (6 de cada jogo) na base criada para esta dissertação. A coluna da esquerda exibe a imagem original enquanto que a coluna da direita exibe a imagem segmentada. As regiões das imagens segmentadas demarcadas com intensidade preta representam áreas do gramado, enquanto que as regiões com intensidade branca correspondem às regiões de não grama.

Tabela A.1 – Imagens originais da base e imagens segmentadas pela abordagem proposta.


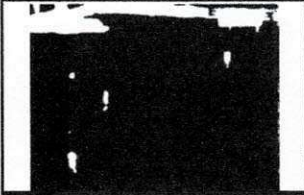

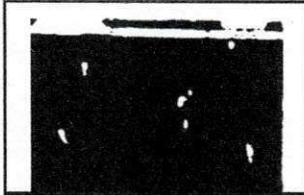

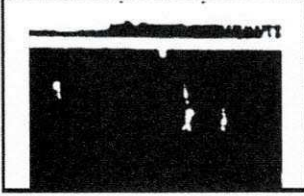



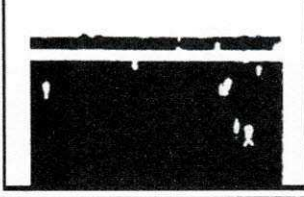



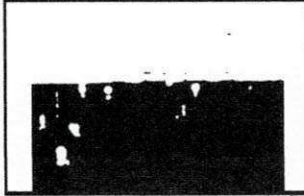

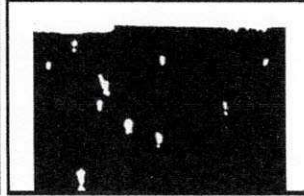

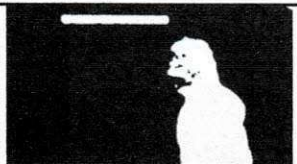



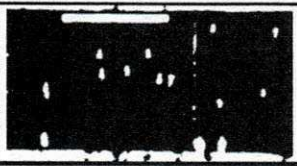





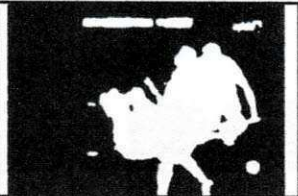





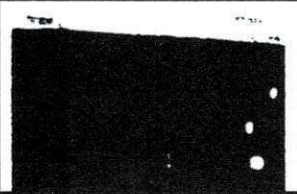

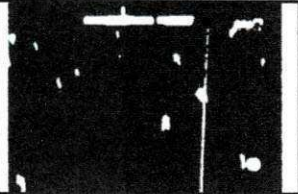

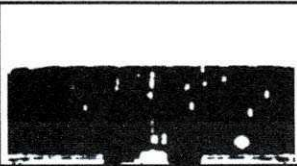



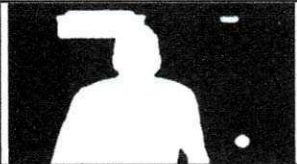





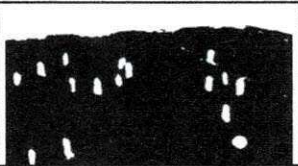
Imagem Original	Imagem Segmentada	Imagem Original	Imagem Segmentada
Jogo 01 – Alemanha Vs Austrália			
			
			
			
Jogo 02 – Argentina Vs Alemanha			
			
Imagem Original	Imagem Segmentada	Imagem Original	Imagem Segmentada

Imagem Original	Imagem Segmentada	Imagem Original	Imagem Segmentada
			
			
Jogo 03 – Argentina Vs Nigéria			
			
			
			
Jogo 04 – Barcelona Vs Manchester			
			
			
			
Imagem Original	Imagem Segmentada	Imagem Original	Imagem Segmentada

Imagem Original	Imagem Segmentada	Imagem Original	Imagem Segmentada
Jogo 05 – Brasil Vs Holanda			
			
			
			
Jogo 06 – Brasil Vs Portugal			
			
			
			
Jogo 07 – Brasil Vs Chile			
			
Imagem Original	Imagem Segmentada	Imagem Original	Imagem Segmentada

Imagem Original	Imagem Segmentada	Imagem Original	Imagem Segmentada
Jogo 08 – Camaroes Vs Holanda			
Jogo 09 – Espanha Vs Portugal			
Imagem Original	Imagem Segmentada	Imagem Original	Imagem Segmentada

Imagem Original	Imagem Segmentada	Imagem Original	Imagem Segmentada
			
Jogo 10 – Franca Vs Mexico			
			
			
			
Jogo 11 – Internacional Vs Corinthians			
			
			
			
Jogo 12 – Manchester Vs Chelsea			
			
Imagem Original	Imagem Segmentada	Imagem Original	Imagem Segmentada

Imagem Original	Imagem Segmentada	Imagem Original	Imagem Segmentada
			
			
Jogo 13 - Palmeiras Vs Americana			
			
			
			
Jogo 14 - Paraguai Vs Espanha			
			
			
			
Imagem Original	Imagem Segmentada	Imagem Original	Imagem Segmentada


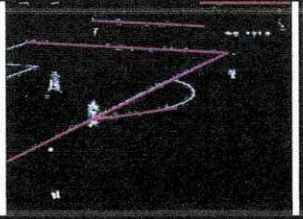



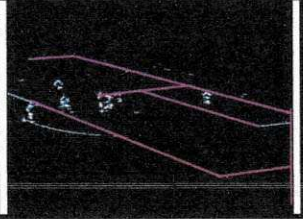

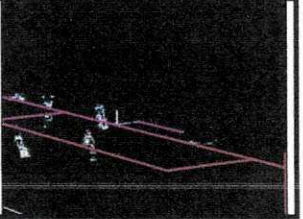

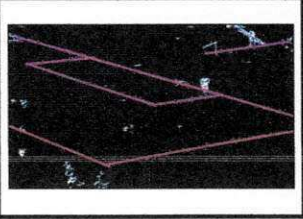

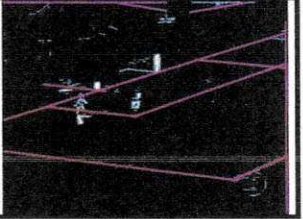
Apêndice B

Resultados da Segmentação das Linhas do Campo


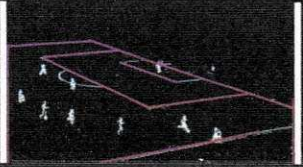
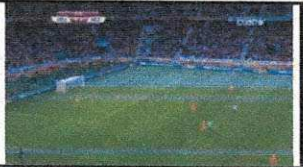
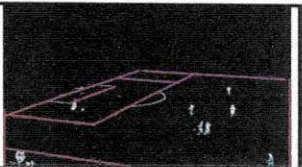



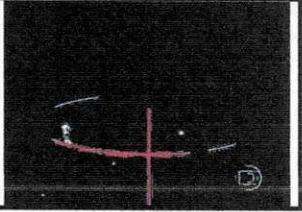



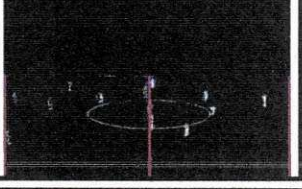




Neste apêndice, apresentam-se exemplos de imagens após a aplicação da segmentação das linhas do campo descrita na Seção 4.2. Na

Tabela B.1, apresenta-se o resultado final de algumas imagens existentes na base criada para esta dissertação. As retas destacadas em vermelho são resultado da aplicação do algoritmo desenvolvido.

Tabela B.1 – Segmentação das Linhas do Campo por meio da abordagem proposta.

Nº	Imagem Original	Imagem Segmentada	Imagem Original	Imagem Segmentada
01				
02				
03				
Nº	Imagem Original	Imagem Segmentada	Imagem Original	Imagem Segmentada

Nº	Imagem Original	Imagem Segmentada	Imagem Original	Imagem Segmentada
04				
05				
06				
07				
08				
09				
10				
11				
Nº	Imagem Original	Imagem Segmentada	Imagem Original	Imagem Segmentada

Nº	Imagem Original	Imagem Segmentada	Imagem Original	Imagem Segmentada
12				
13				
14				
15				
Nº	Imagem Original	Imagem Segmentada	Imagem Original	Imagem Segmentada





Apêndice C

Resultados da Classificação de Tomadas

Apesar de ter-se obtido um excelente resultado, a classificação de tomadas descrita na Seção 4.3 ainda apresenta algumas imagens que não foram corretamente classificadas. Na Tabela C.1, exibem-se algumas destas imagens em que o rótulo Positivos Falsos representa as imagens da classe longa distância classificadas de forma incorreta. Na Tabela C.2 o rótulo Negativos Falsos representa as imagens da classe outras classificadas de forma incorreta.

A falha na classificação de tomadas exibida nas imagens presentes possivelmente foi ocasionada por falhas na segmentação dos blobs como se pode verificar nas imagens (b) e (c) que não obtiveram valor 0, indício que a imagem é pertencente à classe outras. Por outro lado, os erros exibidos na Tabela C.2 podem ser provenientes do grande percentual de grama presente nas imagens (a) e (b) e, ainda, do grande número de *blobs* - (c) e (d); características de imagens pertencentes à classe longa distância.

Tabela C.1 – Imagens “Positivos Falsos” na Classificação de tomadas.

Positivos Falsos			
(a)	(b)	(c)	(d)
			

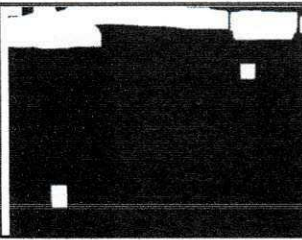
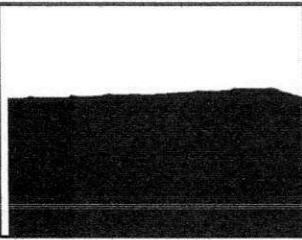
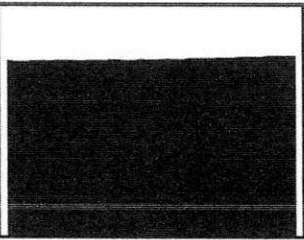
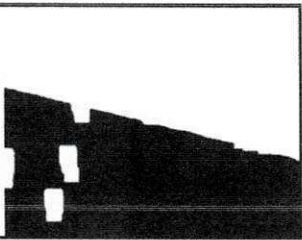
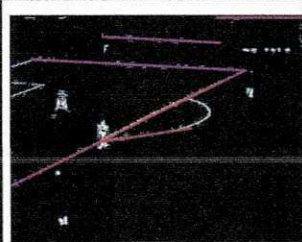

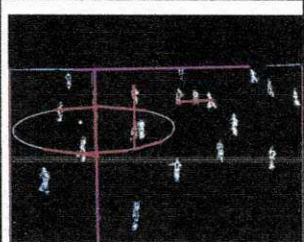
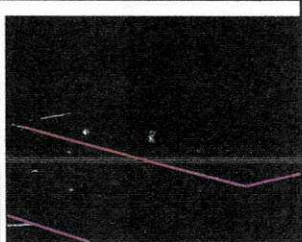




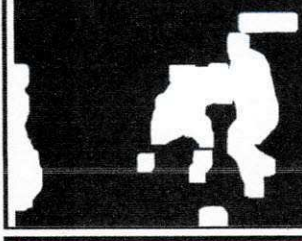
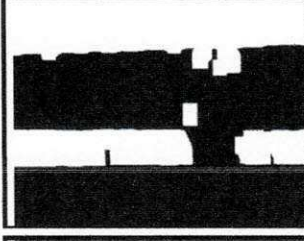
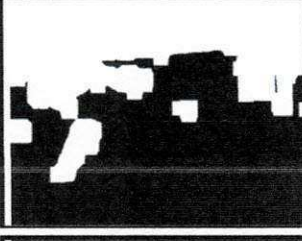
			
			
Percentual de grama = 89,1%	Percentual de grama = 62,86%	Percentual de grama = 77,16%	Percentual de grama = 47,36%
Número de Linhas = 5	Número de Linhas = 2	Número de Linhas = 11	Número de Linhas = 3
Número de Blobs Filtrados = 2	Número de Blobs Filtrados = 0	Número de Blobs Filtrados = 0	Número de Blobs Filtrados = 2

Tabela C.2 - Imagens "Negativos Falsos" na Classificação de tomadas.

Negativos Falsos			
(a)	(b)	(c)	(d)
			
			
Percentual de grama = 89,6%	Percentual de grama = 81,06%	Percentual de grama = 64,56%	Percentual de grama = 59,03%
Número de Linhas = 3	Número de Linhas = 2	Número de Linhas = 11	Número de Linhas = 0
Número de Blobs Filtrados = 4	Número de Blobs Filtrados = 4	Número de Blobs Filtrados = 5	Número de Blobs Filtrados = 5

