

UNIVERSIDADE ANHEMBI MORUMBI
BRUNO NICOLETTI SIQUEIRA
MÁRCIO DOMINGUES MARINS

UMA PROPOSTA DE ANÁLISE DIGITAL DO EXERCÍCIO DE AGACHAMENTO

São Paulo
2009

BRUNO NICOLETTI SIQUEIRA
MÁRCIO DOMINGUES MARINS

UMA PROPOSTA DE ANÁLISE DIGITAL DO EXERCÍCIO DE AGACHAMENTO

Trabalho de Conclusão de Curso apresentado como exigência parcial para a obtenção de título de Bacharel do Curso de Ciências da Computação na Universidade Anhembi Morumbi.

Orientadora: Prof^ª. Marisa Ortegoza da Cunha

São Paulo
2009

UNIVERSIDADE ANHEMBI MORUMBI

BRUNO NICOLETTI SIQUEIRA

MÁRCIO DOMINGUES MARINS

UMA PROPOSTA DE ANÁLISE DIGITAL DO EXERCÍCIO DE AGACHAMENTO

Trabalho de Conclusão de Curso apresentado como exigência parcial para a obtenção de título de Bacharel do Curso de Ciências da Computação na Universidade Anhembi Morumbi.

Aprovado em

Prof^ª. Marisa Ortegoza da Cunha
Universidade Anhembi Morumbi

Prof. Luciano Freire
Universidade Anhembi Morumbi

Prof^ª. Regiane Aparecida Marucci
Universidade Anhembi Morumbi

AGRADECIMENTOS

Agradecemos aos nossos familiares e verdadeiros amigos que nos apoiaram e compreenderam durante tantos finais de semanas, ou mesmo outros dias ocupados e trabalhando para concluirmos nossa graduação.

Agradecemos também aos nossos mestres, por toda a paciência e dedicação de ensinar cada matéria ao longo de todos estes quatro anos. Particularmente, agradecemos nossa orientadora, Marisa Ortegoza, que esteve conosco e ajudou a conduzirmos sem desvios este trabalho até o final.

RESUMO

O objetivo deste trabalho é desenvolver um sistema para avaliação biomecânica baseado na aquisição de imagens que auxilie no treinamento de esportistas de forma relativamente simples e com baixo custo. A análise visa à identificação e extração de características a fim de se estabelecer dados e parâmetros para quantificação de fatores que possam melhorar a execução do movimento do exercício de agachamento. Para isto é necessário a compreensão do movimento do agachamento de um ponto de vista biomecânico, tais como angulação dos joelhos em relação ao solo, posição dos quadris, cabeça e pés. Este sistema envolve a utilização de uma câmera digital, um computador e um sistema capaz de interpretar as imagens obtidas.

Palavras chave: Captura de Movimento, Visão Computacional, Processamento de Imagens, Agachamento.

ABSTRACT

The main goal of this work is to develop a system for bio-mechanic analysis based on image acquisition for improving the training of sportiest in a simple and low cost way. The analysis aims to identify and extract the characteristics in a way of establishing data and parameters for improving the execution of the squat exercise's movement. In order to accomplish that, it's mandatory the comprehension of the squat movement in a bio-mechanic point of view, such as position of the knees, waist, head and feet. This system involves the utilization of a digital camera, a computer and a system capable of understanding the obtained images.

Key words: Motion Capture, Computer Vision, Image Processing, Squat.

SUMÁRIO

1	INTRODUÇÃO	11
1.1	OBJETIVOS	11
1.2	JUSTIFICATIVAS	12
1.3	ABRANGÊNCIA	12
1.4	ESTRUTURA DO TRABALHO	12
2	O EXERCÍCIO E A TECNOLOGIA	14
2.1	O MOVIMENTO.....	14
2.1.1	Classificação do Movimento	14
2.2	O AGACHAMENTO	15
2.2.1	Os Pés	16
2.2.2	Os Joelhos	16
2.2.3	As Costas	17
2.2.4	Os Fêmures	17
2.2.5	Conclusões sobre o Agachamento	17
2.3	VISÃO COMPUTACIONAL.....	18
2.3.1	Visão Humana	19
2.3.2	Aquisição da Imagem Digital	20
2.3.3	O que é um Pixel	21
2.3.4	Arquivo de Imagem Digital	21
2.3.5	Ilusão de Movimento	21
2.3.6	Conclusões sobre a Visão Computacional	22
2.4	ANÁLISE DO MOVIMENTO HUMANO COM VÍDEO	22
2.5	CAPTURE DE MOVIMENTOS	23

2.5.1	Captura de Movimento com Uso de Marcadores	24
2.5.2	Captura de Movimento Sem Marcadores	25
2.6	RECONHECIMENTO DAS FORMAS HUMANAS	25
2.6.1	Remoção do Fundo da Cena	25
2.6.2	Intensidade dos Pixels	26
2.6.3	Intensity Histogram	26
2.6.4	Thresholding	27
3	METODOLOGIA	28
3.1	AMBIENTE.....	28
3.1.1	Iluminação	28
3.1.2	Plano de Fundo	28
3.1.3	Esportista	29
3.1.4	Restrições	29
3.2	HARDWARE	29
3.2.1	Câmera	29
3.3	SOFTWARE.....	30
3.3.1	Processamento do Vídeo	31
3.3.1.1	Segmentação dos Vídeos em Frames.....	31
3.3.2	Processamento de Imagens	32
3.3.2.1	Pré-Processamento.....	32
3.3.2.1.1	<i>Filtro para Remoção do Plano de Fundo</i>	33
3.3.2.1.2	<i>Filtro de Thresholding</i>	33
3.3.2.1.3	<i>Filtro para Remoção de Pixels Isolados</i>	33
3.3.2.1.4	<i>Filtro para Redução de Área</i>	34

3.3.2.2	Extração de Características	36
3.3.2.3	Vetorizando a Imagem Humana	36
4	DESENVOLVIMENTO DO SISTEMA	38
4.1	DIAGRAMA DE CASOS DE USO	38
4.1.1	Manter Vídeos	39
4.1.2	Submeter Novo Vídeo	39
4.1.3	Excluir Vídeo	39
4.1.4	Exibir Vídeo	39
4.1.5	Elaboração do Diagrama de Casos de Uso	40
4.2	ROBÔ PROCESSADOR	40
4.2.1	Captação do Vídeo	41
4.2.2	Fluxo de Processamento do Vídeo	41
4.2.3	Fluxo de Processamento dos Frames	42
4.2.3.1	Pré-Processamento de Imagens	43
4.2.3.2	Extração de Características	44
4.2.3.2.1	<i>Extração dos dados do Crânio</i>	<i>45</i>
4.2.3.2.2	<i>Extração dos dados da Cintura</i>	<i>46</i>
4.2.3.2.3	<i>Extração dos dados do Joelho</i>	<i>46</i>
4.2.3.2.4	<i>Extração dos dados do Pé</i>	<i>47</i>
4.2.4	Construção de um Modelo Vetorizado	47
4.3	ANÁLISE DA CENA	47
4.3.1	Postura dos Joelhos	47
4.3.2	Postura da Cintura	48
4.3.3	Postura da Coluna	48

5	CONCLUSÕES E TRABALHOS FUTUROS	49
	REFERÊNCIAS BIBLIOGRÁFICAS	51

1 INTRODUÇÃO

Processamento de imagens é uma das áreas de estudo, das ciências da computação, que se encontra em constante evolução devido à viabilização de inúmeras aplicações relacionadas à análise automatizada de imagens e o aprimoramento de imagens.

Uma das primeiras aplicações conhecidas da tecnologia de aprimoramento de imagens digital remonta ao início do século XXI, onde foi empregada na codificação de imagens para o envio através de cabos submarinos entre a Inglaterra e os Estados Unidos da América para que então os dados fossem decodificados do outro lado e então interpretados como imagem.

Mas o grande impulso para a área de Processamento de Imagens viria cerca de três décadas mais tarde, com o advento dos primeiros computadores digitais de grande porte e o início do programa espacial norte-americano. O uso de técnicas computacionais de aprimoramento de imagens teve início no Jet Propulsion Laboratory (Pasadena, Califórnia - EUA) em 1964, quando imagens da lua transmitidas por uma sonda Ranger eram processadas por computador para corrigir vários tipos de distorções inerentes à câmera de TV acoplada à sonda. Estas técnicas serviram de base para métodos aprimorados que realçavam a restauração de imagens de outros programas espaciais posteriores, como as expedições tripuladas da série Apollo, por exemplo (Filho e Neto, p. 23, 1999).

Desde o seu início até os dias de hoje, o processamento de imagens vem crescendo de maneira exponencial, abrangendo cada vez mais segmentos de mercado, e tem sido aplicado para resolver problemas como a alimentação de sistemas a partir de imagens digitalizadas de formulários preenchidos à mão ou na contagem de elementos presentes em imagens digitais, obtidas com o uso de microscópios, na recuperação de imagens danificadas, na identificação biométrica e também na visão computacional, que vem impulsionando o desenvolvimento da robótica.

1.1 OBJETIVOS

Este trabalho tem por objetivo desenvolver um sistema para análise e avaliação do exercício de agachamento, através da captura de imagens, sem que nenhum equipamento seja acoplado ao corpo dos esportistas e utilizando apenas uma câmera filmadora digital de uso doméstico. A proposta é inserir esta solução em um sistema web.

1.2 JUSTIFICATIVAS

Existe hoje uma carência de métodos automatizados de análise de desempenho na área de fisiologia; os estudos de captura de movimento já estão sendo empregados nesta área, porém o alto custo dos equipamentos envolvidos dificulta sua utilização em massa. Este trabalho visa promover o estudo das técnicas de captura de movimento sem o uso de marcadores, que podem diminuir os custos de implantação e manutenção deste tipo de ferramenta, tornando a computação uma ferramenta de uso viável no treinamento esportivo.

1.3 ABRANGÊNCIA

O sistema elaborado deve ser capaz de reconhecer um exercício de agachamento a partir da captura de um vídeo, em um formato digital, seguindo uma série de regras definidas ao longo deste trabalho. Para os propósitos visados, a captura deve ser realizada apenas em um plano, ou seja, com a utilização de apenas uma câmera (2 Dimensões), e o movimento deve ser capturado sem a utilização de marcadores anexados ao corpo do esportista. Este sistema deve contar com uma interface gráfica para a entrada do vídeo e outra para a visualização do modelo vetorial do esportista executando o movimento, exibindo dados sobre a execução para cada um dos quadros selecionados e, ao término apresentar um resumo com informações e sugestões de melhorias na execução.

1.4 ESTRUTURA DO TRABALHO

Ao longo deste trabalho, são descritas todas as etapas do desenvolvimento de um sistema que contemple as necessidades descritas no item 1.1. Assim, este trabalho foi estruturado da seguinte maneira:

Será abordado no 2º capítulo o estudo da literatura, buscando o embasamento científico sobre o movimento humano e sua contextualização no estudo do agachamento: a visão computacional e a visão humana, de forma a buscar o entendimento do que é uma imagem para o computador. Esse conhecimento, que torna possível a compreensão do relacionamento entre a análise do exercício físico e a tecnologia computacional, possibilitará analisar técnicas de captura de movimento em um computador, incluindo conceitos técnicos já bem estabelecidos.

Após estes conceitos, serão avaliados no 3º capítulo, os pré-requisitos, desde o ambiente apropriado para poder acompanhar o exercício, até os equipamentos que podem ser utilizados para a captura do movimento. Será arquitetada uma solução plausível nesse contexto. Nesse capítulo, serão mostradas técnicas desenvolvidas para viabilizar a extração de características nos parâmetros determinados no objetivo.

Vistos esses tópicos, toda informação necessária para gerar uma proposta de implementação terá sido suprida. Esta é apresentada no 4º capítulo, descrevendo desde casos de uso até as API's necessárias e os algoritmos especialistas que farão a extração das características do corpo humano.

Finalmente, as conclusões são apresentadas e este trabalho deixará uma abordagem e uma prospecção para futuros trabalhos, correlacionados nesta área.

2 O EXERCÍCIO E A TECNOLOGIA

De acordo com Delavier (2000, p. 78), o desenvolvimento da musculatura provou-se de grande valia na obtenção de bons resultados, fazendo com que a musculação fosse inserida no sistema de treinamento de esportistas.

Para o aprimoramento das qualidades e desenvolvimento de volume muscular, é necessário conhecer a anatomia e a fisiologia do corpo humano - sem estes conhecimentos, os resultados podem não ser os esperados e há possibilidade de ocorrência de lesões, impossibilitando o atleta de competir. O agachamento é considerado por muitos como o "rei dos exercícios" e, segundo Gentil (2009), é um dos exercícios mais completos que podem ser realizados dentro das academias, pois envolve um elevado número de articulações e músculos. Por outro lado, muitos praticantes e pesquisadores condenam este exercício, afirmando que ele promove lesões e danos musculares. Para evitar este tipo de lesão, os esportistas são orientados e acompanhados por profissionais conhecedores das técnicas, que os observam durante a execução do agachamento.

Remondino (2004, p.1-5) afirma que a demanda de modelos humanos digitais está crescendo drasticamente, o que estimulou o desenvolvimento de diversas técnicas de captura de movimentos. Algumas destas técnicas envolvendo o uso de marcadores especiais já são empregadas na reabilitação de esportistas lesionados. O uso destes marcadores, porém, inviabiliza sua aplicação no acompanhamento diário de esportistas, seja pelo seu custo elevado, como também pela sua complexidade de utilização.

2.1 O MOVIMENTO

Uma partícula no espaço pode ser identificada com um ponto representado pelas coordenadas (x, y, z) . Quando se estuda o movimento de uma partícula, estamos, na verdade, estudando a variação das coordenadas desta partícula em relação ao tempo.

2.1.1 Classificação do Movimento

O movimento é classificado de acordo com a natureza do objeto observado. Segundo Tavares (2004, p. 2-3), o movimento rígido é aquele em que, dada uma variação de tempo, a

distância entre dois pontos do objeto é sempre preservada. Já o articulado, é o movimento de um conjunto de elementos, envolvendo partes rígidas conectadas por ligações não rígidas.

Estas classificações serão levadas em consideração na proposta deste trabalho. Analisando os tipos de movimentos descritos, podemos considerar que o corpo humano apresenta um movimento articulado. Observando o modelo humano na Figura 1, fica fácil identificar os vértices e arestas que o representam.

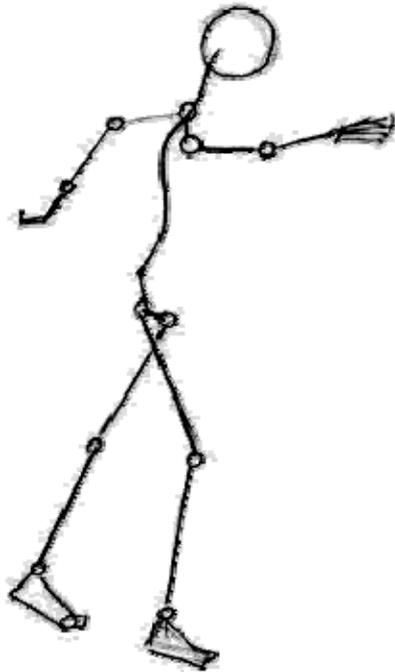


Figura 1 – Identificação de arestas e vértices no modelo humano. Fonte: O Autor (2009).

2.2 O AGACHAMENTO

O agachamento é um exercício de musculação que tem como objetivo promover o ganho de força e massa nos músculos dos membros inferiores do corpo.

O movimento é executado dobrando-se as pernas nos joelhos e quadris, abaixando o tronco entre as pernas, depois invertendo a direção e levantando-se. Durante o processo, o tronco deve permanecer ereto. Sidney (2003) afirma que o esportista atinge a amplitude máxima do exercício quando seus fêmures ficam horizontais. Outra visão a ser considerada é a de que os joelhos não devem ultrapassar a ponta dos dedos dos pés, fazendo com que o quadril mantenha-se para trás. Analisando a Figura 2 fica simples visualizar a correta execução.

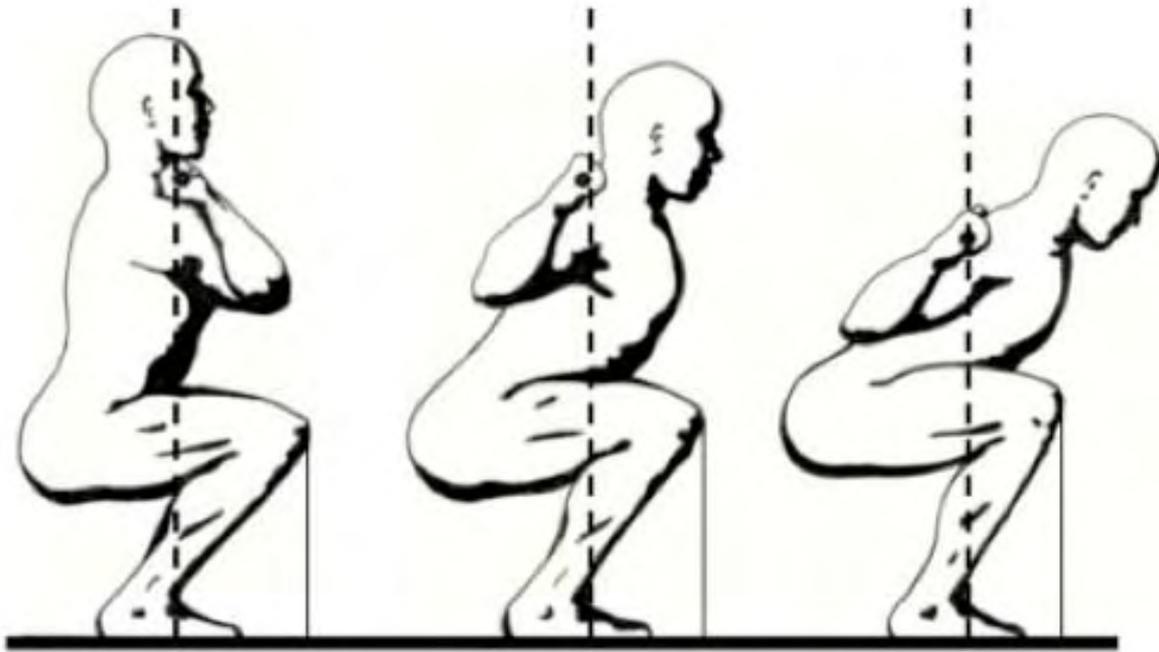


Figura 2 – Visão lateral do agachamento. Fonte: Rippetoe (2007).

Para avaliar um esportista durante um treinamento de agachamentos e identificar possíveis falhas na execução, é importante primeiramente avaliar individualmente o papel de cada uma das partes do corpo humano envolvidas no movimento e identificar o comportamento esperado e o que pode caracterizar uma falha de execução.

2.2.1 Os Pés

Os pés possuem um papel importante no agachamento: eles transferem toda a força dos quadríceps para o solo. Durante a execução, o esportista deve permanecer com os pés completamente em contato com o solo, colocando o peso nos calcanhares.

Para validar se os pés tocam por completo o solo durante todo o período da execução, o esportista deve ser observado lateralmente. O observador deve identificar se a ponta dos pés e os calcanhares sofrem alteração verticalmente; caso ocorra algum tipo de elevação vertical em um desses pontos, uma falha na execução é identificada.

2.2.2 Os Joelhos

Oliveira (2002) afirma que o agachamento é perigoso, pois ao flexionar o joelho em ângulos maiores que 90° , aumenta-se perigosamente a tensão na patela. Para reduzir a tensão

no joelho gerada pelas cargas, o esportista deve jogar o quadril para trás, mantendo os joelhos para trás da ponta dos pés.

Para a melhor observação, o observador deve se posicionar de maneira que ele tenha uma visão lateral do esportista e então identificar a ponta do pé e o joelho. Ao longo do exercício, ele deve comparar a posição deles horizontalmente e verificar se o joelho passa a linha da ponta do pé.

2.2.3 As Costas

Outra premissa do agachamento, e talvez a mais importante, é a de que as costas devem se manter eretas durante toda a execução. De acordo com os estudos de Oliveira (2008), o agachamento só é prejudicial à coluna nos casos de prática indiscriminada. Recomenda-se que o esportista mantenha as costas eretas para evitar estas lesões.

Ao manter as costas eretas, o esportista fixa a distância da sua cabeça ao seu quadril; caso algum tipo de curvatura ocorra esta distância é comprometida, o que caracteriza um movimento rígido da coluna. Qualquer variação na distância entre a cabeça e o quadril pode ser considerada uma curvatura da região lombar, caracterizando uma falha de execução.

2.2.4 Os Fêmures

"A amplitude do agachamento é muito importante", "grande parte dos estudos e recomendações limitando o movimento, se referem ao agachamento paralelo, realizado até que as coxas fiquem paralelas ao solo o que gera amplitudes maiores que 90 graus de flexão dos joelhos" (Gentil, 2009).

A posição dos fêmures em relação ao solo pode ser validada através da verificação do ângulo formado entre uma reta traçada a partir da posição dos joelhos até a base do quadril e o solo a partir de uma observação lateral. Qualquer ângulo negativo formado entre o fêmur e o solo é considerado falha de execução.

2.2.5 Conclusões sobre o Agachamento

Com base nas informações levantadas sobre a correta análise de um exercício de agachamento, o menor grupo de pontos necessários incluem a ponta do pé, o calcanhar, o

joelho, o quadril e a cabeça. A Figura 3 exibe a completa identificação destes pontos e mostra por que o melhor ponto de observação é definido em uma visão lateral do esportista.

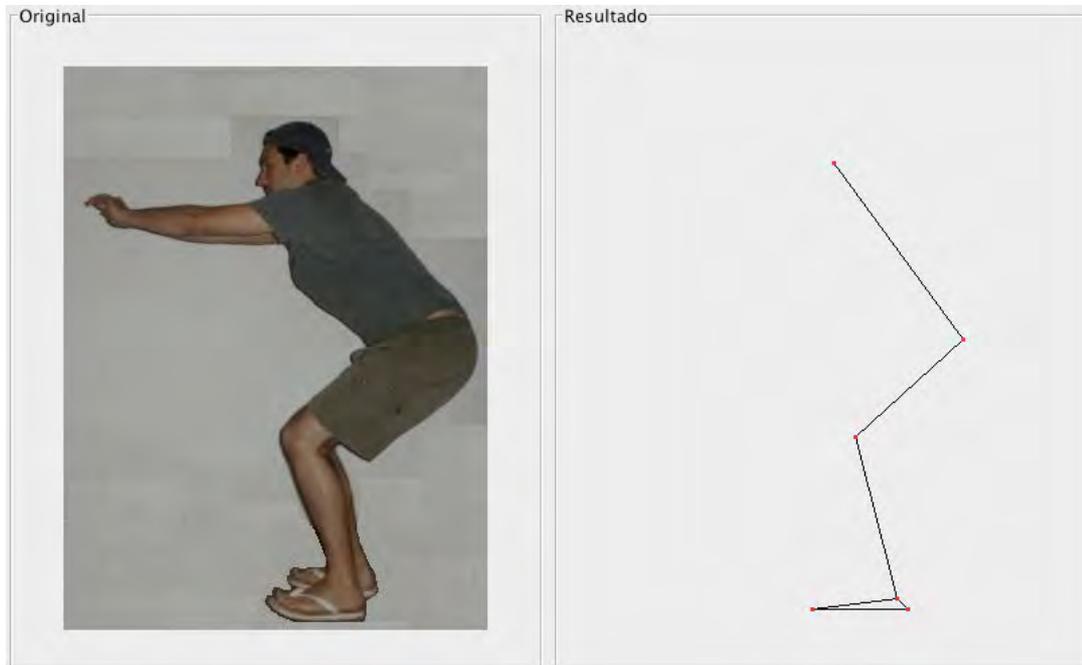


Figura 3 – Visão lateral do agachamento Fonte: O autor (2009).

2.3 VISÃO COMPUTACIONAL

Visão computacional é a ciência e tecnologia envolvidas no estudo de técnicas que permitem aos computadores “compreender” o conteúdo de uma imagem, onde “compreender” deve ser entendido como: Extrair dados suficientes para que um determinado problema possa ser resolvido.

De acordo com os autores (Gomes e Velho, p. 2 e 3, 1994), existem quatro áreas que atuam de um modo unificado e cooperativo, conforme o diagrama da Figura 4. Este trabalho atua tanto na Visão Computacional quanto no Processamento de Imagens, visto que a imagem recebida deverá ser tratada através de filtros de forma a diminuir a complexidade da imagem e só então serem aplicadas técnicas específicas para a extração de dados ou características.

A visão computacional se assemelha muito à visão humana em diversos aspectos e para entender o seu funcionamento é necessária uma compreensão básica do modo como o sistema de visão humana funciona.



Figura 4 – Computação Gráfica e Áreas Correlatas. Fonte: Gomes e Velho (1994)

2.3.1 Visão Humana

O olho é o órgão responsável pela captura das imagens no corpo humano. Na Figura 5 podemos observar um corte lateral do olho. Basicamente, o processo para formação das imagens se dá da seguinte maneira. A luz do objeto observado incide através da córnea e se projeta na retina.

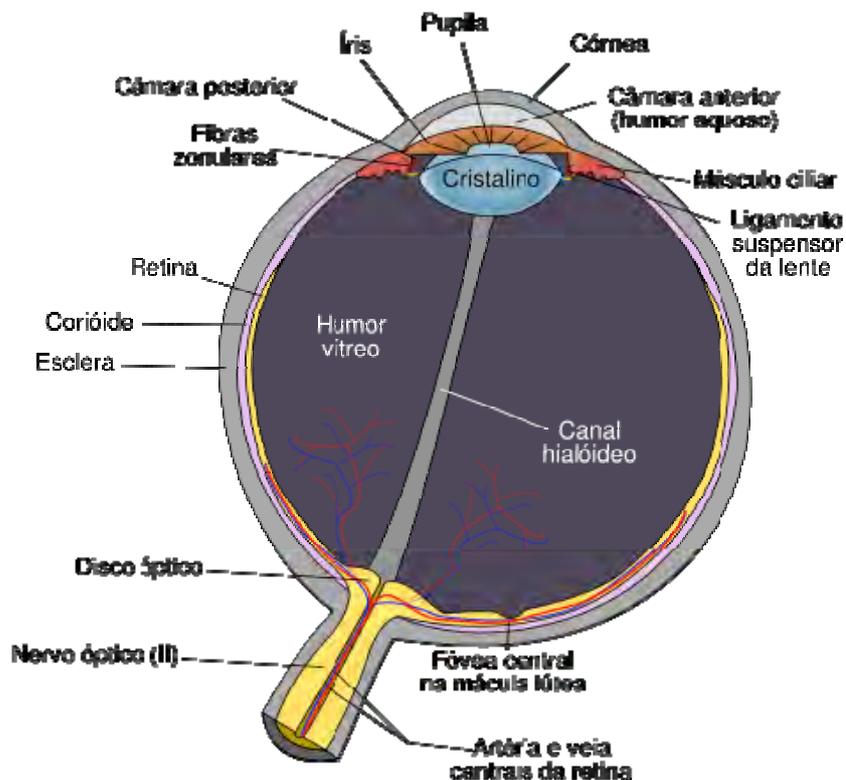


Figura 5 – Elementos do olho humano

A retina possui milhares de células divididas em dois tipos: os bastonetes e os cones, ambos sensíveis à luz. Quando a luz é projetada sobre estas células, estas são estimuladas e, como consequência, geram impulsos elétricos. Quanto mais luz, maior será a estimulação e maior será o impulso gerado. Estes pulsos gerados são enviados para o nervo ótico e de lá são direcionados para o cérebro, onde são interpretados.

2.3.2 Aquisição da Imagem Digital

No corpo humano os olhos são responsáveis pela aquisição da imagem, convertendo luz em impulsos elétricos. Já em um sistema de visão computacional, o papel dos olhos é realizado pelas câmeras digitais que convertem luz em elétrons e, posteriormente, em bits.

As câmeras digitais formam a imagem de maneira semelhante à do olho humano. A luz incide por um ponto - as lentes - e é projetada em uma área coberta por células que reagem à luz. No caso das câmeras digitais, essas áreas preenchidas por células fotos sensíveis são chamadas de Dispositivos de Cargas Acopladas ou apenas CCD (*Charge-Coupled Devices*).

As células fotos sensíveis presentes no CCD, assim como os bastonetes e cones no olho humano, reagem com a luz, gerando eletricidade. Esta eletricidade gerada tem uma variação diretamente proporcional à quantidade de luz recebida pela célula. Esses dados capturados pelo CCD são imediatamente convertidos em uma matriz de bytes, sendo que cada uma de suas células gera a informação referente a um pixel da figura montada, conforme vemos na Figura 6.

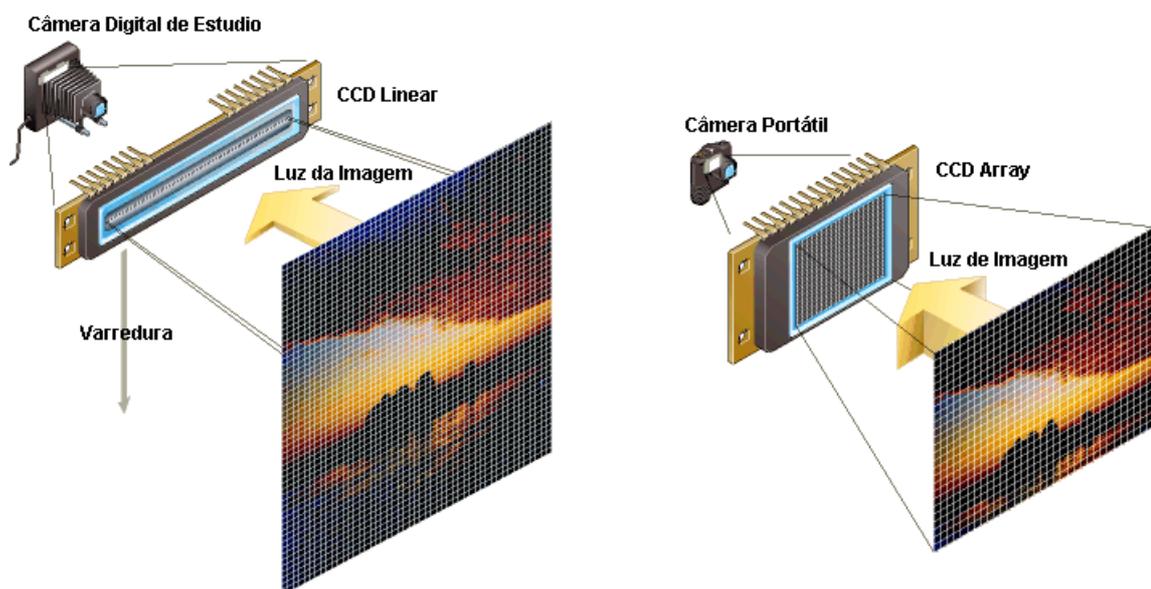


Figura 6 – Elementos do olho humano. Fonte: Cappelaro (2009).

2.3.3 O que é um Pixel

Um pixel é o menor elemento em uma imagem digital a que se pode atribuir uma cor. De uma forma mais simples, um pixel é o menor ponto que forma uma imagem digital, sendo que o conjunto de milhares de pixels forma a imagem inteira. Aos pixels são atribuídos valores numéricos que representam a sua cor no modelo RGB (Red, Green, Blue).

2.3.4 Arquivo de Imagem Digital

Um arquivo de imagem digital é formado por um conjunto de pixels. Estes arquivos possuem um tamanho pré-definido pela quantidade de pixels contidos. Os pixels são organizados em uma matriz onde a largura da imagem é expressa pela quantidade de pixels por coluna e a altura, pela quantidade de linhas.

Ao final, o arquivo por si só é apenas uma representação abstrata: uma vez exibida e capturada pelo sistema de visão humano, o cérebro é o responsável por transformar esses dados em uma imagem.

2.3.5 Ilusão de Movimento

Filme é uma coleção de imagens capturadas e exibidas em seqüência com um intervalo pré-determinado entre elas. Cada uma dessas imagens recebe o nome de quadro. O cérebro humano pode ser iludido se a quantidade de quadros exibidos por segundo for alta o suficiente, e interpretar essa seqüência de imagens como um movimento contínuo. Por este motivo as câmeras de vídeo digitais trabalham capturando uma série de imagens em seqüência mantendo sempre uma taxa de quadros por segundo estática.

Para que a captura do vídeo seja suficiente para uma futura análise biomecânica do agachamento, devemos levar em consideração dois fatores: a quantidade de quadros necessária para uma correta análise, e o tempo médio de execução do exercício.

O tempo de execução de uma repetição do exercício pode variar de acordo com o tipo de treinamento, mas podemos considerar que uma repetição dificilmente será mais rápida do que 1 segundo e mais lenta do que 5 segundos.

Visando a atender um dos objetivos deste trabalho, que é propor uma solução relativamente barata para captura de imagens, vamos considerar a utilização de câmeras de

vídeo digital para o uso amador. Estas câmeras seguem em média um padrão de captura de 24 quadros por segundo.

Com 24 quadros por segundo numa execução rápida (1 segundo) serão capturados aproximadamente 24 quadros, já em uma execução lenta, que a pessoa leve até 4 segundos, serão gerados aproximadamente 100 quadros. Sendo assim podemos afirmar que uma câmera de uso amador é suficiente para a correta captura do movimento de agachamento.

2.3.6 Conclusões sobre a Visão Computacional

Tais estudos mostram que a tecnologia de hardwares envolvida na captura de imagens e vídeos, já está bem estabelecida. Tal mecanismo imita o funcionamento do olho humano, trazendo para o âmbito mecatrônico uma linguagem apropriada e bem inerente à tecnologia computacional.

2.4 ANÁLISE DO MOVIMENTO HUMANO COM VÍDEO

Eadweard Muybridg foi um pioneiro nos estudos do movimento humano com vídeo, com a intenção de demonstrar que um cavalo, ao galopar, em um determinado momento da ação, retirava as quatro patas do chão.

Muybridg então desenvolveu um sistema onde 24 câmeras escuras foram posicionadas lado a lado e foram conectadas a um mecanismo de disparo ativado pelo próprio cavalo, que ao atingir uma determinada posição, acionava a câmera que, por sua vez, registrava uma fotografia daquele momento. Este disparo era repetido por cada uma das câmeras, cada uma delas registrando o cavalo em um determinado momento. Se compararmos aos filmes atuais cada fotografia seria o equivalente a um quadro; se exibirmos estes quadros em sequência, temos um filme.

Seu projeto foi um sucesso. Ele conseguiu provar sua teoria de que as quatro patas deixavam de tocar o solo em um determinado momento, como visto logo no segundo quadro da Figura 7. Seus estudos, realizados em 1872, são utilizados até hoje como referência para análise de movimento através de imagens.

Como podemos observar, as imagens obtidas por Muybridg ressaltam o cavalo do fundo, para que fique mais fácil analisarmos o movimento sem que ocorra uma falha na interpretação por conta de objetos não relevantes.

Uma das maiores dificuldades em reconhecer objetos não-rígidos, como o corpo humano, é o próprio ambiente em que ele está contido. Dado que o único objeto de relevância para os estudos do movimento humano é o próprio corpo humano, podemos considerar qualquer outro tipo de objeto que apareça na imagem como sendo sujeira ou ruído e, por isto, devem ser descartados.

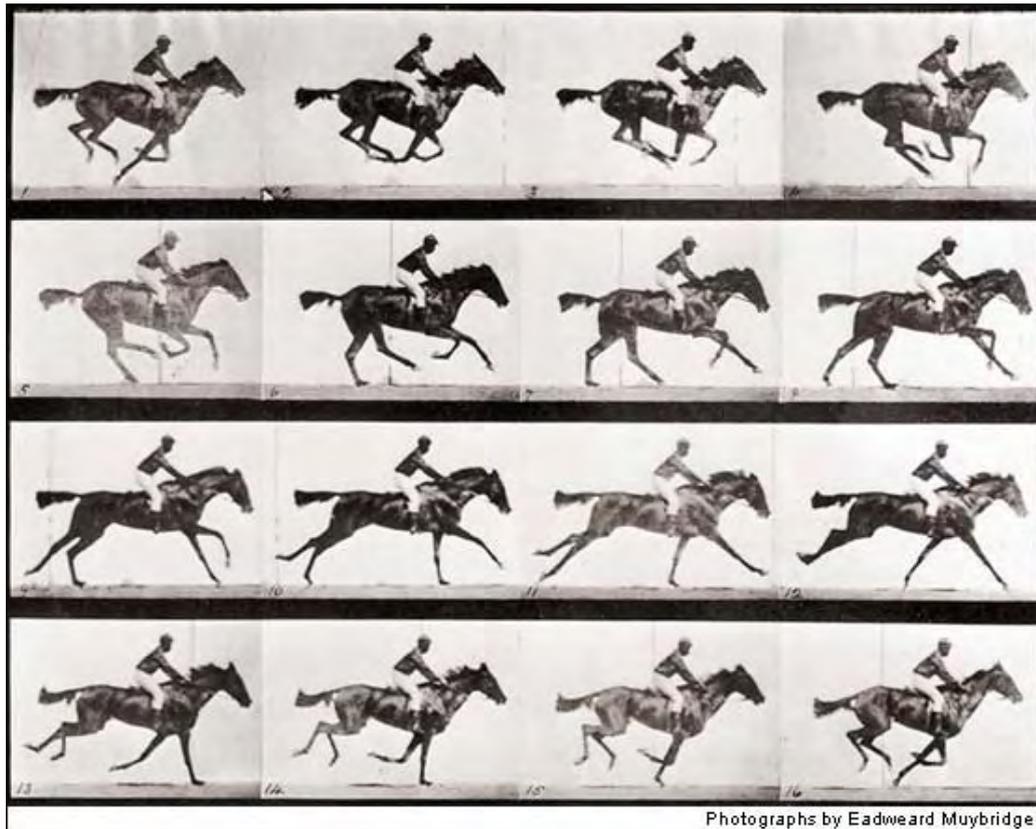


Figura 7 – Análise do movimento a partir da técnica de Muybridg. Fonte: NMeM (2009).

Para que se torne viável o desenvolvimento de um algoritmo capaz de reconhecer as formas humanas, algumas técnicas foram desenvolvidas para simplificar a imagem original, ressaltando apenas o objeto de desejo.

2.5 CAPTURA DE MOVIMENTOS

Captura de movimentos é um termo utilizado para descrever o processo de reproduzir um movimento em um modelo digital.

Na área de vigilância e segurança, a captura de movimento vem sendo amplamente utilizada para identificar pessoas e prevenir ou identificar crimes. A captura de movimentos

também está relacionada com a criação de aplicações para controle de funcionalidades, utilizado hoje como interface de jogos, ambientes virtuais, animações, entre outras. Outra área de utilização é a de análise, onde a captura de movimentos é empregada na reabilitação de pacientes com diagnósticos ortopédicos ou ainda no auxílio de esportistas no sentido de melhorarem seus desempenhos.

2.5.1 Captura de Movimento com Uso de Marcadores

Conforme se pode ver na Figura 8, o objeto observado recebe pequenos dispositivos que são posicionados ao longo do seu corpo. Estes marcadores podem ser construídos com materiais que refletem a luz de volta ou que possuam luminosidade própria. Estes pontos se tornam brilhantes e se destacam na imagem capturada, com a aplicação de filtros, como *Thresholding*, que isolam os pontos mais brilhantes do restante da imagem.



Figura 8 – Uso de marcadores na captura do movimento.

Os marcadores luminosos simplificam drasticamente o sistema de captura do movimento, uma vez que os pontos de observação já estão destacados. Por outro lado, tornam sua captura mais complicada, dado que a pessoa que executa o movimento a ser observado, precisa vestir um traje negro, com vários pontos luminosos acoplados. Este traje, apesar de facilitar o desenvolvimento da captura, é um fator extremamente limitador.

2.5.2 Captura de Movimento Sem Marcadores

Esta técnica é bem pouco difundida. Seu uso baseia-se em softwares que são capazes de encontrar padrões de formas e distinguir tais padrões em uma figura, ou mesmo em um vídeo.

Seria o mesmo princípio de uma rede neural, onde é dado um objeto de entrada para aprendizado e depois o processamento se incumbe de distinguir tal padrão no objeto que lhe é passado. No contexto deste trabalho, esse objeto poderia ser um vídeo. Porém, visto a complexidade do corpo humano e seus membros, não seria possível, desta forma, alcançar os objetivos. Assim, serão utilizadas técnicas especializadas para tais análises, que serão apresentadas no capítulo 3.

2.6 RECONHECIMENTO DAS FORMAS HUMANAS

A primeira etapa para qualquer tipo de processamento de imagens é justamente a simplificação da imagem original, de maneira que a imagem resultante contenha apenas os objetos de relevância para o problema proposto.

Conforme afirmado anteriormente, o corpo humano é o único objeto a ser estudado para uma completa análise do movimento de agachamento. Sendo assim, para simplificar o problema e facilitar o processamento da imagem, a primeira etapa seria justamente a extração da silhueta humana. Para isto, pode-se basear em algumas técnicas já desenvolvidas.

2.6.1 Remoção do Fundo da Cena

De acordo com Davis (2009), a visão humana é caracterizada por associar uma cor constante a um objeto, mesmo que este sofra variações de luminosidade, o que podemos chamar de fidelidade a cor. Esta fidelidade garante a classificação dos objetos.

A técnica de remoção de fundo, também conhecida como *Background Subtraction* consiste em subtrair uma imagem de outra imagem utilizada como referência.

Se utilizarmos como imagem de referência uma imagem onde o objeto a ser identificado não apareça, neste caso o corpo humano, teremos uma imagem contendo apenas o plano de fundo. Ao selecionar um quadro onde o objeto apareça, basta subtrair o plano inicial e o que teremos é apenas a silhueta humana. Podemos ver esse teste quadro a quadro demonstrado na Figura 9.

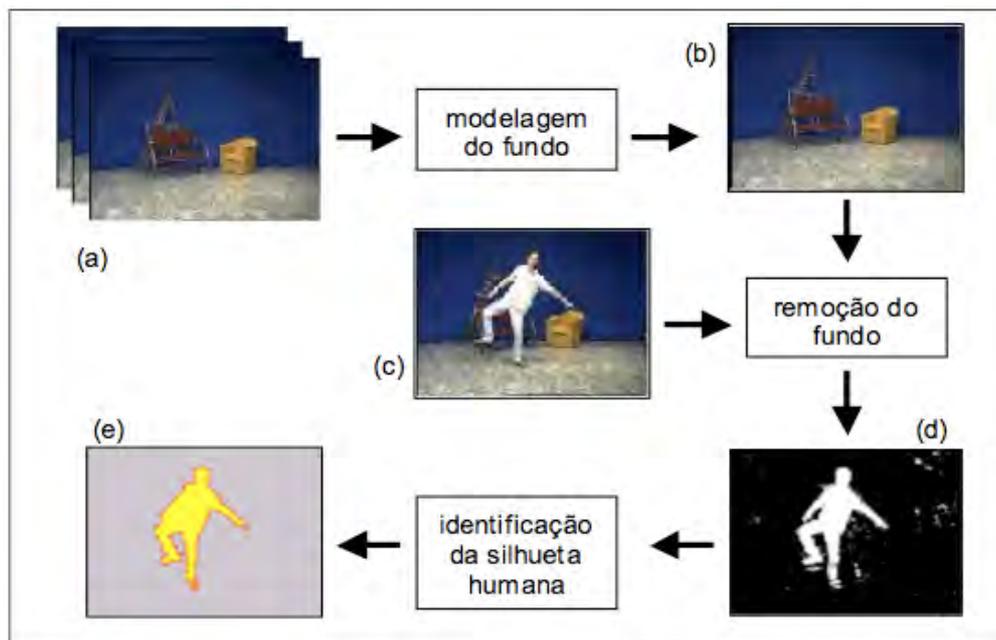


Figura 9 – Etapas para remoção do plano de fundo

A imagem gerada através da subtração dos quadros mostra os pixels classificados como não pertencentes ao plano de fundo, porém nem sempre todos os pixels selecionados fazem parte da figura humana. Estes pixels são chamados ruídos da imagem. Apesar de normalmente aparecerem espalhados e em pequenas quantidades, eles podem atrapalhar os futuros cálculos.

2.6.2 Intensidade dos Pixels

A compreensão de como os pixels são representados é importante para que o processamento de uma imagem seja viável, dado que o pixel é a menor parte de uma imagem digital. Cada um dos pixels guarda, em si, uma informação de cor; no caso das imagens em preto e branco, ele guarda apenas a informação de brilho. Um valor de 0 a 255 é atribuído para cada um destes pixels, sendo 0 o branco absoluto, 255 o preto absoluto e os valores intermediários representando as tonalidades de cinza.

Baseando-se nestas informações sobre os pixels, pode-se extrair diversos tipos de informações sobre a imagem analisada.

2.6.3 Intensity Histogram

Visto na Figura 10, o *intensity histogram*, é um gráfico que associa o número de pixels ao nível de intensidade que lhes é atribuído. Esse gráfico permite a seleção de uma ou mais

regiões de acordo com a intensidade destes pixels.

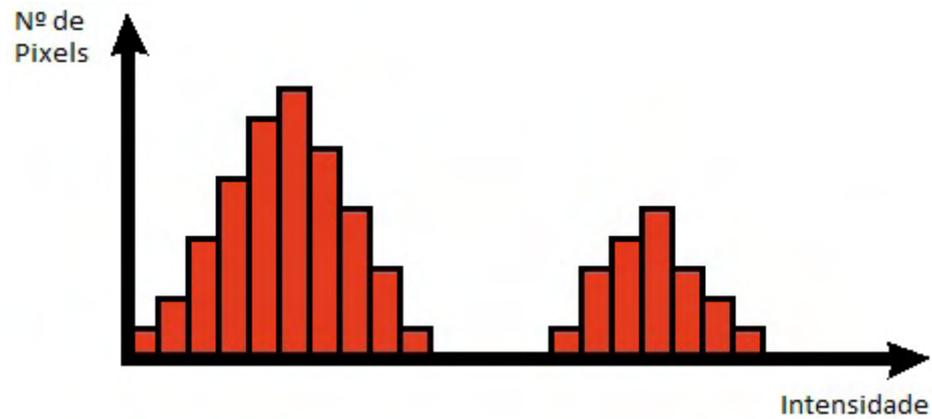


Figura 10 – Gráfico *Intensity Histogram*. Fonte: Fisher (2009).

2.6.4 Thresholding

Thresholding é uma técnica de segmentação de imagens em preto e branco, usualmente utilizada na separação de regiões da imagem.

Esta técnica se baseia no estudo do *Intensity Histogram*. Um corte é realizado separando aqueles pixels que apresentam um nível de intensidade superior e inferior a um determinado valor, de modo que a imagem final perde as informações de níveis de cinza, transformando-se em uma matriz de pixels brancos e pretos.

Na Figura 11 vemos um gráfico com os níveis de escala das cores demonstrando como ficaria com um corte nos limites e outro bem ao meio.

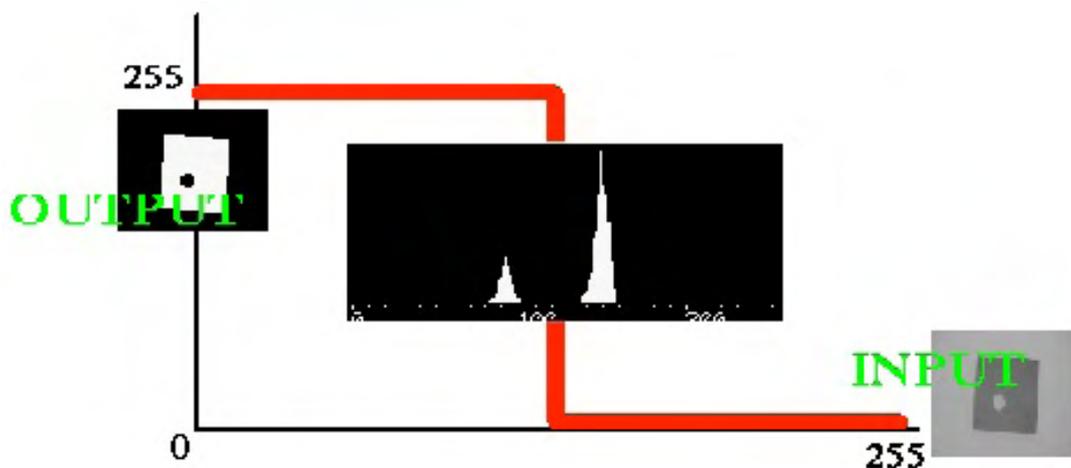


Figura 11 – Demonstração do *Thresholding* no gráfico de *Intensity Histogram*

3 METODOLOGIA

A metodologia aborda a descrição das técnicas e os conceitos exigidos para aplicação em um software capaz de atingir nosso objetivo. São detalhadas as condições ambientais e restrições, seguindo a proposta para a extração de características.

3.1 AMBIENTE

A proposta deste trabalho é a criação de um sistema de baixo custo e de fácil aplicação. O ambiente de captura do movimento deve ser o mais simplificado e menos limitante possível. O grande desafio é eliminar toda a parafernália exigida por softwares de mesmo objetivo, que, em sua maioria, se baseiam em marcadores.

O uso desta tecnologia deve ser viável, mesmo em um ambiente como o de uma academia. Serão analisadas mais à frente restrições, mas da forma mais minimizada dentro do contexto atual.

3.1.1 Iluminação

Para a melhor captura dos dados, é de extrema importância que o ambiente possua uma iluminação adequada. É aconselhável que a luz incida sobre o atleta, de forma a projetar a sombra no solo. Caso a sombra seja projetada na direção oposta do esportista, teremos, na filmagem, além da imagem do esportista, a imagem de sua sombra.

Como dito antes, na seção 2.6.1, a imagem do fundo poderá ser retirada utilizando-se o processo denominado *Background Subtraction*, porém, caso exista a sombra, teremos, uma nova imagem no cenário que pode dificultar a separação exclusiva da silhueta do corpo.

3.1.2 Plano de Fundo

A proposta deste trabalho é facilitar o uso de tal tecnologia e para tanto foram estudadas formas para a maneira mais prática da aplicação com o uso extensivo. Com isso, foram adotadas técnicas que visaram à solução dentro de tais parâmetros.

Visto as solicitações de Iluminação na seção 3.1.1 atendidas, outra preocupação é a de simplificação do plano de fundo. Mesmo com a aplicação de técnicas de remoção do plano de

fundo, recomenda-se que este seja preferencialmente uma superfície lisa e principalmente sem existência de material refletor no ambiente. Não é admitido de forma alguma o uso de um espelho atrás da câmera filmadora. Um material reflexivo poderia alterar o cenário, trazendo qualquer imagem de algo que se movimenta atrás da câmera, alterando assim a primeira imagem que é utilizada como modelo para subtração do fundo.

3.1.3 Esportista

Para evitar a distorção da silhueta humana, recomenda-se que o esportista vista roupas com tecidos finos e justos. A lycra é um material recomendado por possuir todas as características necessárias e manter sua elasticidade, não atrapalhando o esportista durante a execução do movimento.

As técnicas empregadas no reconhecimento da forma humana são especialistas e trabalham através de padrões pré-definidos em software. Caso o atleta não se enquadre no perfil esperado, o reconhecimento é invalidado.

3.1.4 Restrições

Nesta seção serão descritas as restrições encontradas ao longo de testes que serão realizados com o software em construção.

3.2 HARDWARE

A proposta é simplificar ao máximo possível, portanto o hardware requerido é apenas uma câmera de vídeo comum. Esta poderá tanto ser uma câmera de um aparelho de celular, com resolução mínima de 240 x 320 e podendo gravar a 24 fps (frames por segundo), quanto uma sofisticada câmera filmadora.

3.2.1 Câmera

O vídeo obtido pela câmera é a entrada do sistema, que fornece todos os dados necessários para a análise do exercício. Conforme comentado na seção 2.2.5, a visão lateral da execução do exercício proporciona a observação de todos os pontos necessários para uma análise do

movimento de agachamento. Além do posicionamento lateral da câmera e preparação do ambiente conforme orientado na seção 3.1, é de extrema importância que a câmera faça a captura do sujeito por completo, de maneira que a câmera não sofra mudanças de posicionamento durante a captura.

Para isto, recomenda-se a utilização de um suporte fixo, como um tripé, a uma distância fixa, de maneira que o sujeito seja capturado por completo durante todos os momentos da sua execução. O posicionamento da câmera pode ser realizado da seguinte maneira: a câmera deve ficar a uma distância de 2 metros do esportista e a uma altura de 1 metro do solo; caso este posicionamento não seja suficiente para a captura completa do corpo do esportista, adaptações podem ser realizadas.

Além das preocupações com posicionamento, outro ponto de atenção é a de que o exercício seja capturado seguindo uma ordem de atividade.

O vídeo deve ser iniciado com a captura do ambiente vazio, durante um tempo não menor do que 1 segundo, para que seja possível a extração do plano de fundo.

3.3 SOFTWARE

Nesta seção apresenta-se o desenvolvimento de um software capaz de analisar as imagens obtidas seguindo os padrões descritos anteriormente. Este software tem o objetivo de processar um vídeo e retornar um conjunto de dados sobre a execução do movimento de agachamento realizado no vídeo.

O software desenvolvido deve contemplar uma interface web para acesso dos usuários, com as seguintes propriedades:

- a) Interface para cadastro de usuários;
- b) Interface para autenticação;
- c) Interface com a página inicial do usuário autenticado, contendo um resumo do histórico das análises executadas;
- d) Uma interface para o envio de um novo vídeo para análise;
- e) Interface contendo o relatório digital da análise do exercício, com imagens dos pontos críticos apontando as falhas de execução.

O processamento dos vídeos enviados deve ser realizado em um processo separado. Neste processamento, o software deve aplicar conceitos de extração de características executando um conjunto de operações específicas para cada tipo de padrão identificado. O sistema deve

ter a capacidade de analisar as imagens do vídeo, identificando primeiramente os pontos essenciais do corpo do esportista, de criar um modelo vetorial do mesmo, e, a partir dos dados obtidos, identificar a fase do movimento. Cada uma destas fases possui um conjunto de regras que definem um movimento ideal - para cada uma destas fases deve ser criada uma série de operações específicas. Ao final deste processamento, o sistema deve atribuir um valor que define o grau de qualidade da execução. Além disto, o sistema deve fazer críticas e sugestões para o aprimoramento da execução.

O processamento dos vídeos deve respeitar o fluxo de trabalho conforme representado na Figura 12.



Figura 12 – Fluxo de processamento do vídeo. Fonte: O autor (2009).

3.3.1 Processamento do Vídeo

O processamento do vídeo é o início de todo o processo de análise digital. Conforme visto na seção anterior, o sistema web recebe um vídeo para que seja realizado o processamento. Um processo separado, rodando em background, deve observar um diretório, aguardando o aparecimento de um vídeo. O sistema web deve direcionar o arquivo para este diretório.

Uma vez localizado um arquivo para processamento, o sistema deve segmentar o vídeo em uma coleção de quadros. Estes quadros são extraídos de acordo com uma regra de distância entre eles.

3.3.1.1 Segmentação dos Vídeos em Frames

Um filme é uma seqüência de quadros e, para que exista uma ilusão de movimento, o padrão é que sejam exibidos pelo menos 24 quadros por segundo.

Visto que um exercício de agachamento leva em média em torno de 1 a 3 segundos por repetição, podemos então garantir uma média de aproximadamente 30 quadros por repetição. Nem todos os quadros precisam ser analisados, e para garantir um desempenho aceitável do sistema, definimos uma regra onde o sistema deve analisar 1 a cada 3 quadros do filme, o que resulta numa média aproximada de 10 quadros analisados por repetição do movimento.

3.3.2 Processamento de Imagens

O processamento do vídeo faz a extração dos quadros gerando uma coleção de imagens a serem processadas. O processamento das imagens é subdividido em 2 etapas: o pré-processamento e a extração das características.

3.3.2.1 Pré-Processamento

A primeira etapa do processamento é a simplificação da imagem recebida. O objetivo desta etapa é extrair a silhueta humana removendo qualquer outro tipo de objeto da imagem, obtendo assim uma imagem onde o único objeto exibido é o corpo do esportista em preto e branco. Uma série de processamentos é executada na imagem, onde a entrada é sempre a imagem original e a saída é outra imagem, já tratada. Para este tipo de processamento, vamos atribuir o nome ‘filtros’. Conforme podemos observar na figura 13, os ‘filtros’ aplicados são:

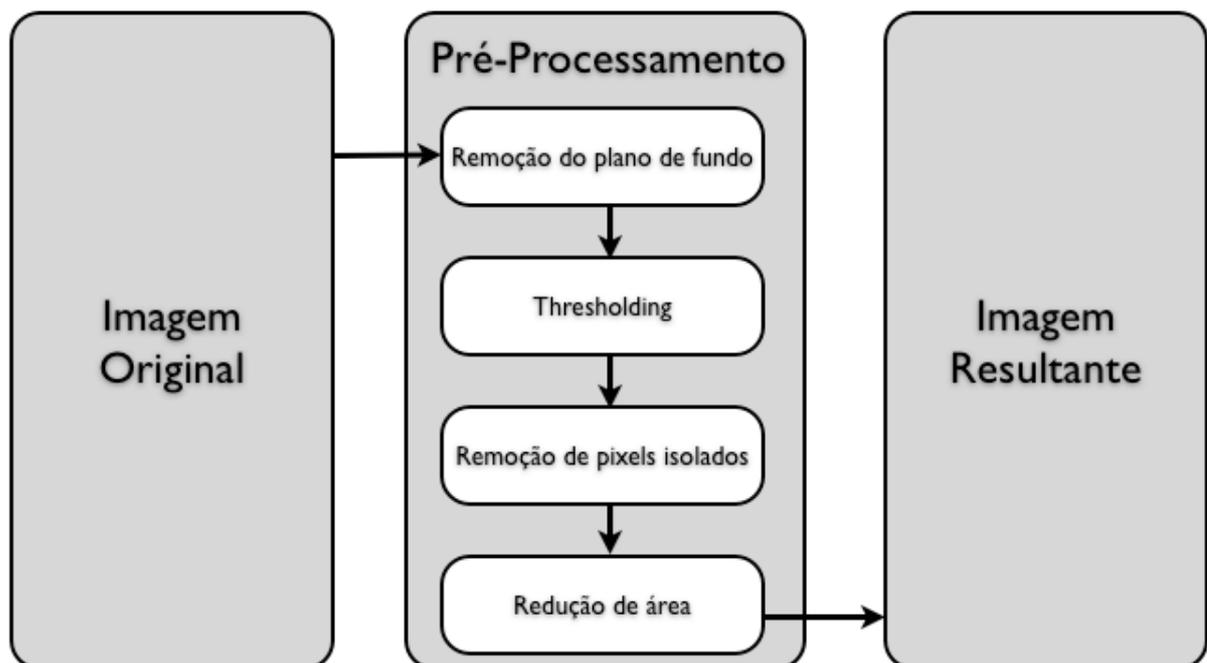


Figura 13 – Sequência de aplicação dos filtros. Fonte: O autor (2009).

3.3.2.1.1 Filtro para Remoção do Plano de Fundo

O sistema deve iniciar a resolução do problema proposto aplicando um filtro de remoção do plano de fundo, como na técnica descrita em 2.6.1, utilizando ‘*Background Subtraction*’

Para que seja realizada a subtração dos quadros posteriores, o primeiro quadro recebido pelo sistema é sempre armazenado em memória, gerando assim uma imagem mais limpa. Ao final deste filtro, é esperado que a silhueta do corpo do atleta se destaque do fundo, porém é esperada ainda a existência de algumas sujeiras em menores proporções.

3.3.2.1.2 Filtro de Thresholding

No item 2.6.4 foi visto que o *Thresholding* é uma técnica para seleção de pixels de acordo com sua intensidade. Este filtro se mostrou de grande utilidade na extração de sujeiras resultantes da remoção do plano de fundo. Ao final da aplicação deste filtro, a imagem resultante pode ainda conter alguns pixels isolados conforme figura 14.

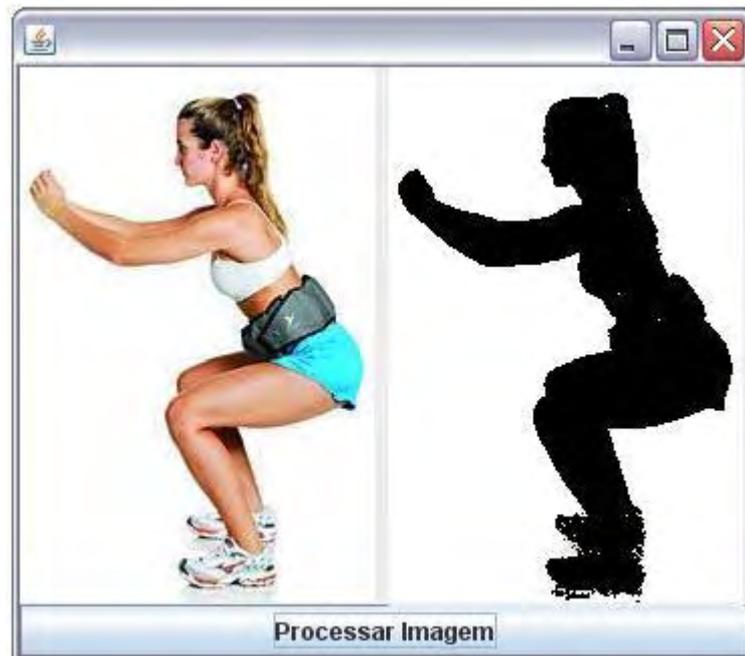


Figura 14 – Aplicação do filtro *thresholding*. Fonte: O autor (2009).

3.3.2.1.3 Filtro para Remoção de Pixels Isolados

Este filtro tem como objetivo a remoção de sujeiras em menor escala, que possam atrapalhar os cálculos seguintes. Uma vez identificado como sendo um pixel isolado,

lembrando que a esta altura a imagem trabalhada está em preto e branco, este tem sua cor original invertida, ou seja, se ele é branco, passa a ser preto na imagem resultante, e vice-versa.

Um pixel isolado pode ser um pixel branco rodeado por pixels pretos, ou o contrário, mas em alguns casos podemos ter mais de um pixel juntos; por este motivo é considerado como isolado todo pixel que contém menos de 4 pixels vizinhos de mesma cor.

Na figura 15 podemos observar alguns pixels marcados em vermelho: são os pixels que seriam considerados como isolados. Como podemos observar, alguns pixels são perdidos, como por exemplo, os dos cantos do quadrado maior na Figura 16. Estas perdas são esperadas e não exercem grande impacto na análise da imagem gerada, uma vez que os pixels que atrapalham os cálculos são aqueles que se misturam à imagem e acabam gerando dados de distância inconsistentes.

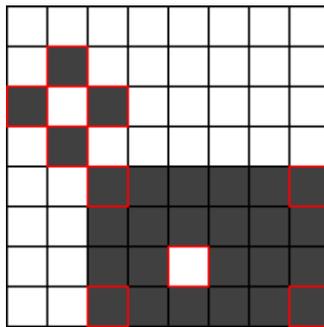


Figura 15 – Pixels isolados. Fonte: O autor (2009).

3.3.2.1.4 Filtro para Redução de Área

Depois da aplicação de todos os filtros anteriores, a imagem resultante esperada deve conter uma área branca com apenas a silhueta do corpo humano em preto. Dependendo da forma como a câmera foi posicionada, a silhueta humana pode preencher uma área grande ou pequena da imagem total. Quanto maior a imagem, maior é a quantidade de pixels que a representa e, conseqüentemente, maior será o tempo de processamento necessário. Para os propósitos deste trabalho, a única área de interesse a ser analisada é a que representa o corpo humano; portanto, aplica-se um último filtro, responsável por identificar e extrair uma imagem com a menor área possível e que contenha todos os pixels pretos.

Este filtro identifica os limites de início e fim para cada um dos eixos e, ao final, faz um recorte da área interna desses limites. Estes limites são identificados da maneira como segue.

A matriz de pixels é percorrida primeiramente da esquerda para a direita, linha a linha. Quando o primeiro pixel preto é localizado, guardamos a sua posição em relação ao eixo Y, definindo assim o limite superior, conforme Figura 16.

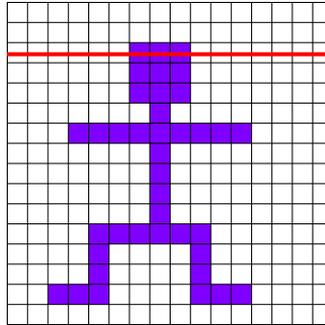


Figura 16 – Identificação do limite superior para corte. Fonte: O autor (2009).

Logo em seguida, a matriz é novamente percorrida, porém desta vez, de cima para baixo, coluna a coluna. Quando o primeiro pixel preto é localizado, guardamos a sua posição em relação ao eixo X, definindo assim o limite esquerdo, conforme Figura 17.

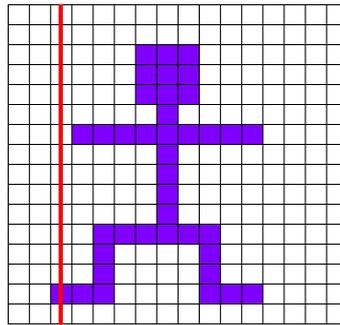


Figura 17 – Identificação do limite esquerdo para corte. Fonte: O autor (2009).

Para localizar o limite inferior e direito, o procedimento é repetido, porém invertendo a ordem de execução, ou seja, para o limite inferior a matriz é percorrida da esquerda para a direita e as linhas são percorridas de baixo para cima. Para o limite direito, ela é percorrida de cima para baixo e as colunas são percorridas da direita para a esquerda, conforme a Figura 18.

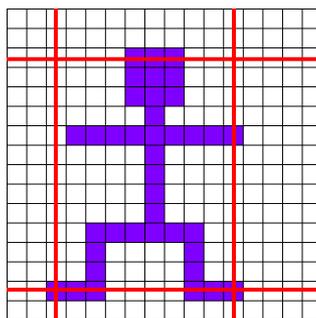


Figura 18 – Identificação da área de corte. Fonte: O autor (2009).

Os limites identificados formam uma área que representa o resultado da redução esperada. O filtro então conclui seu trabalho, fazendo um recorte e gerando como saída esta imagem de menor área.

3.3.2.2 Extração de Características

Nesta etapa, após a imagem já processada, iniciaremos a fase da ‘Visão Computacional’ onde o objetivo é obter informações geométricas para compor nosso modelo 2D (2 dimensões) e extrair os dados necessários para a avaliação do movimento de nosso esportista.

Como o corpo humano tem medidas proporcionais, iremos dividir a imagem primeiramente, em quatro partes iguais, no eixo latitudinal, para análise individual detalhada, conforme ilustra a figura 19.

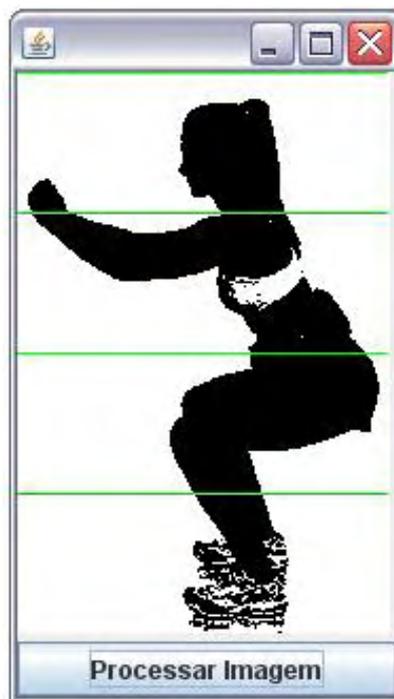


Figura 19 – Análise fracionada em 4 partes iguais. Fonte: O autor (2009).

3.3.2.3 Vetorizando a Imagem Humana

Pensando em cada parte separadamente, pode-se gerar algoritmos capazes de percorrer a matriz de pixels que temos em mãos agora, nossa imagem tratada e fracionada, e buscar formas padrões.

De cima para baixo temos:

Na 1º Parte: a Cabeça;

Entre 2º e a 3º Parte: o Quadril;

Na 3º Parte: os Joelhos;

Na 4º Parte: as pontas dos pés e os calcanhares;

Ainda é possível determinar outras frações para melhorar a precisão da divisão e mesmo analisar cada parte isoladamente, cada uma com uma fração mais específica. O importante é entender que, uma vez com a imagem simplificada e fracionada, pode-se gerar algoritmos capazes de percorrer entre os pixels ‘pretos’ e fixar vértices que serão base de nosso esqueleto, os quais, ligados, formarão nosso modelo 2D. Tais algoritmos serão abordados no capítulo 4.

4 DESENVOLVIMENTO DO SISTEMA

Neste capítulo será mostrado como fazer uso de todas as técnicas propostas e implementar o sistema proposto anteriormente.

Este se trata de um sistema especialista, com a parte mais importante sendo executada por scripts em background. Foi definida para a interação com o usuário a tecnologia web. O usuário deste sistema deverá preencher um cadastro; após este passo, ganhará as credenciais que irão gerenciar suas informações pessoais de treinos e performance.

4.1 DIAGRAMA DE CASOS DE USO

Para o atendimento completo das funcionalidades, três interfaces com o usuário serão suficientes; interface de *'Login'*, de Cadastro e uma que chamaremos de “Manter Vídeos”, que dará acesso a outras funcionalidades.

A interface inicial é a de *'Login'*. Esta é bem simples e seu intuito é disponibilizar os campos para preenchimento do apelido ou *'user'* e para a senha. Haverá também um mecanismo para recuperação da senha: em caso de esquecimento, será feito o *'reset'* da senha gerando uma nova, e enviando-a para o email do usuário.

No caso de novo usuário, ele encontrará um link na interface de *'login'* que irá redirecioná-lo para a interface de cadastro, na qual deverá preencher os seguintes campos:

Apelido ou *'user'*: Informação que sempre será solicitada no acesso ao sistema na interface de *'Login'*.

Nome e sobrenome: Informações que serão utilizadas na personalização do sistema, exibindo estes dados em interfaces de acesso e na exibição de relatórios.

'E-Mail': Este dado é de suma importância, pois será o meio de comunicação do sistema com o usuário para recuperação de senhas e confirmação de cadastro.

Senha: Serão dois campos na interface de cadastro que previnem erros de digitação, durante o cadastramento. Este dado sempre será solicitado na interface de *'Login'*.

Uma vez autenticado, o usuário irá navegar pela interface Manter Vídeos. Esta é a interface onde ele encontrará todas as opções para o gerenciamento dos dados que o sistema administra.

4.1.1 Manter Vídeos

Esta interface gerencia todos os acessos e baseia-se em uma lista que contém as informações de todos os treinos já postados. Selecionando uma opção em um registro desta lista, o usuário pode ser remetido à interface de Submissão, Exclusão ou Exibição de Vídeos. Esta lista é ordenada por data de submissão, onde as mais antigas vão para o fim da lista.

4.1.2 Submeter Novo Vídeo

Neste momento, o usuário fará o *'upload'* do vídeo, disponibilizando-o para a rotina de processamento, executada em *'background'*. O resultado só poderá ser acessado na lista de vídeos; caso haja algum problema, será descrito um *'log'* de erro na Lista de Vídeos.

A rotina que executa em *'background'* será bem detalhada no capítulo que descreve o Robô Processador. Serão descritas todas as fases de processamento deste vídeo e a forma como são extraídas os dados e armazenados.

4.1.3 Excluir Vídeo

Qualquer registro da lista poderá ser escolhido para exclusão. Caso o usuário não deseje mais ver determinado treino, poderá marcar os registros e eliminar uma seqüência de um treino selecionada.

4.1.4 Exibir Vídeo

Uma vez solicitada a exibição de um registro, o usuário será redirecionado a uma interface com todos os detalhes apurados de um vídeo submetido. Esta interface lista os valores apurados em cada execução do movimento. As colunas irão quantificar os pontos críticos do movimento, definidos no capítulo de Agachamento:

- a) Grau da sobre coxa em relação ao solo;
- b) O quanto os joelhos passaram em relação à ponta dos pés;
- c) Qual o grau da base dos pés em relação ao solo, lembrando que os calcanhares devem ficar junto ao solo;

d) O quanto as costas destoam de uma linha reta. Será definida uma unidade de medida para mensurar esta análise;

e) Um valor inteiro, entre 5 níveis possíveis, determinando a qualidade do movimento. Este valor definirá uma nota final para cada execução da seqüência postada.

4.1.5 Elaboração do Diagrama de Casos de Uso

É importante saber que cada vídeo define uma série. Por exemplo, em um treino o atleta executa 3 séries de 10 movimentos. Logo, este postará 3 vídeos, um de cada série. O resultado de um item detalhado da lista determina um vídeo postado, com os valores extraídos da amplitude completa de um movimento.

No fim, haverá um totalizador de eficiência, fruto do cálculo da média aritmética de todos os exercícios da seqüência executada.

Na figura 20, pode-se visualizar o Diagrama de Caso de Uso que exhibe todas as ações que o usuário poderá efetuar no sistema. Este diagrama trata o processamento do vídeo como uma rotina de um elemento externo, por ser uma rotina executada em *'background'*, sem qualquer contato com o usuário após a submissão do vídeo, gerida apenas pelo Robô Processador.

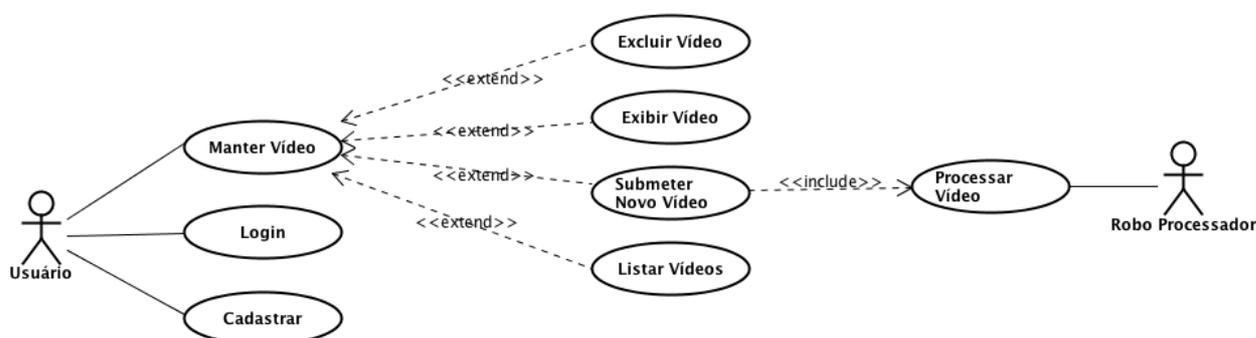


Figura 20 – Diagrama de Caso de Uso do sistema. Fonte: O autor (2009).

4.2 ROBÔ PROCESSADOR

Como dito anteriormente, o processamento do vídeo inicia-se após a sua submissão. Esta submissão é feita pelo *'up load'* de um arquivo de vídeo de extensões como avi, gsm, mov ou wav. Somente serão aceitos tais tipos de mídias.

O Robô Processador tem várias fases até finalizar a extração, todas já citadas no capítulo 3. Iremos detalhar nos próximos itens todas as fases necessárias para transformar o vídeo em dados.

4.2.1 Captação do Vídeo

O Robô Processador verifica, em um tempo pré-determinado, a existência de arquivos no diretório destinado ao *'up load'*. Uma vez encontrado um arquivo, ele irá colocar o vídeo em uma fila de Processamento; depois de obtidos os resultados do Processamento do Vídeo, ele irá eliminar o arquivo, limpando o diretório.

Para a manipulação do vídeo, utilizamos a 'API' JMF (Java *Media Framework*) e Marvin Framework. Estas API's permitem receber um vídeo no ambiente Java, instanciá-lo como Objeto e aplicar os algoritmos propostos para trabalhar o vídeo.

4.2.2 Fluxo de Processamento do Vídeo

A próxima fase do Robô Processador é fragmentar o vídeo em frames; neste momento, a mídia recebida é instanciada como um Objeto `javax.media.player` no código. O algoritmo irá checar se é um arquivo de vídeo válido e, caso não seja, deverá ser criado um log. Sendo verdadeiro, este Objeto é submetido ao processo de extração de imagens. Esse processamento separa um em cada três frames do Objeto e os guarda em um *Array*. No próximo passo, este *Array* de *frames* é submetido a outro processador, dessa vez o Processador de Imagens.

A Figura 20 define o diagrama deste processamento, auxiliando no seu entendimento.

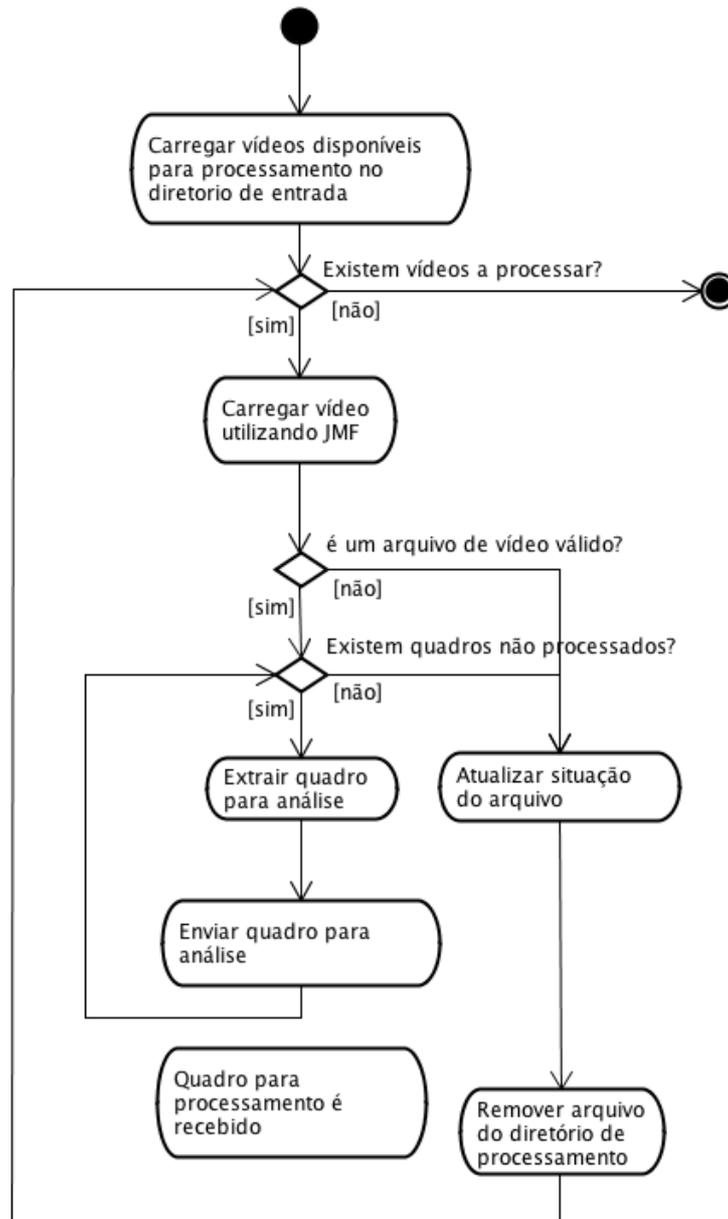


Figura 20 – Fluxo de Processamento do Vídeo. Fonte: O autor (2009).

4.2.3 Fluxo de Processamento dos Frames

Como descrito no capítulo 3, o Processamento de Imagens é dividido em duas fases: o Pré-Processamento de Imagens e a Extração de Características. Ou seja, este passo tem como objetivo duas tarefas:

Pré-Processamento de Imagens – prepara a imagem a ser analisada. Momento em que são aplicados filtros para tratar a imagem, visando a diminuição da complexidade, limpeza de ruídos e padronização de proporcionalidade.

Extração de Características – momento em que são aplicados os algoritmos especialistas propostos no capítulo 3, buscando padrões de formas e definindo pontos de articulação; estes que serão base para formar o modelo vetorial da forma humana.

4.2.3.1 Pré-Processamento de Imagens

Como demonstra o fluxograma da Figura 21, este processador recebe um *Array* de Imagens que, submetido a um *looping*, irá fazer o tratamento das Imagens. Uma vez iniciado o *looping*, será separada a primeira imagem para tomarmos como base da técnica do *background subtraction*. Neste momento, serão utilizados recursos já existentes na API Marvin Framework e JAI (Java *Advanced Image*).

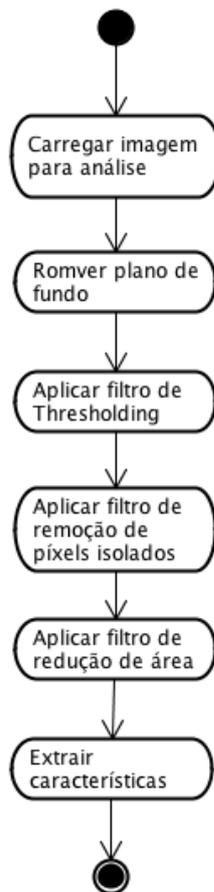


Figura 21 – Fluxo de Pré-Processamento das Imagens. Fonte: O autor (2009).

Os próximos passos serão aplicados da segunda à última imagem. Remoção do plano de fundo, para se ater somente na pessoa, que é a imagem que interessa.

Filtro de *Thresholding*, como já mencionado, define a nossa imagem em apenas duas cores, preto e branco, além de limpar alguns ruídos. A remoção de pixels isolados foi uma técnica que propusemos para refinar o resultado do filtro de *Thresholding*.

O filtro de redução de área, proposto no capítulo 3, nos devolve uma imagem com dimensões dentro dos padrões humanos, dando-nos condições de esperar proporcionalidade. Os algoritmos de extração de características propostos usam o resultado deste filtro; afinal, segmentamos a imagem em valores percentuais, considerando proporções de figuras humanas e esperando encontrar formas conhecidas em cada quadrante separadamente.

4.2.3.2 Extração de Características

A extração das características é o processo responsável por transformar uma imagem em dados, no caso, valores que representam os pontos capturados nos eixos X e Y.

Este processo de extração é realizado de maneira específica por ponto, ou seja, para cada característica. Por exemplo, na análise da cabeça, existirá um algoritmo responsável pela extração de suas características. Descreveremos, a seguir, as técnicas utilizadas para a extração de cada um dos pontos vistos anteriormente como desejáveis para análise do exercício de agachamento.

As características humanas desejadas para extração e posterior vetorização são: Cabeça, Cintura, Joelho e Pé. A Figura 21 demonstra o fluxo de processamento desta etapa.

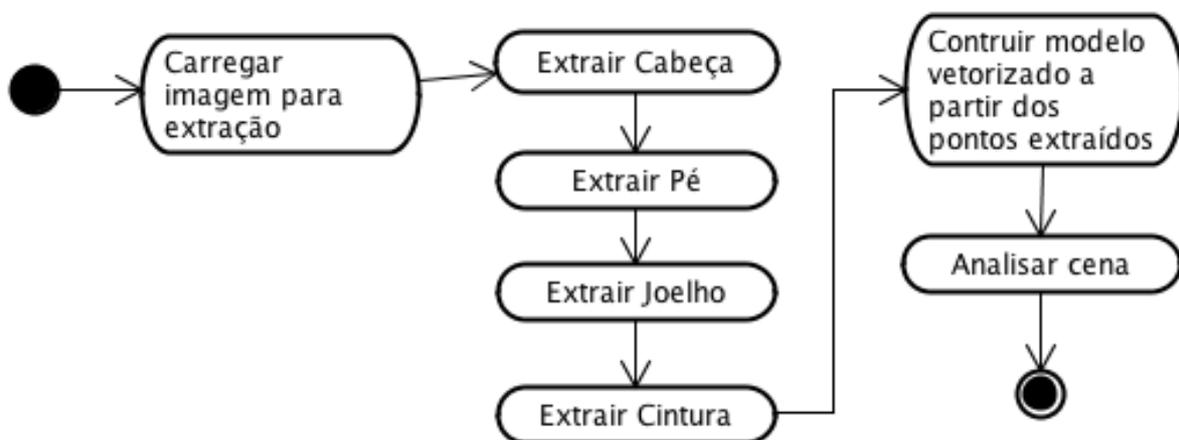


Figura 21 – Fluxo de Extração de Características. Fonte: O autor (2009).

Todos os extratores construídos seguem um fluxo básico de vida, onde primeiramente o extrator deve selecionar uma área para recorte. Esta etapa foi adicionada para simplificar o

processamento da próxima etapa, responsável por transformar a imagem em valores numéricos.

A seleção da área de recorte consiste basicamente em selecionar um pedaço da imagem original onde os pontos desejáveis estarão contidos. Conforme vimos anteriormente, o exercício de agachamento será capturado na posição lateral e o filtro de redução de área irá reduzir a imagem original ao mínimo possível. Desta maneira, podemos garantir, por exemplo, que a cabeça sempre estará na parte superior da imagem, os pés na parte inferior e a cintura e joelhos no centro.

A redução da área a ser processada de acordo com os interesses do extrator facilita o processamento futuro; contudo, implica a alteração da posição dos pontos num plano cartesiano.

Após o recorte da imagem, o plano muda e sua posição acompanha esta mudança. Para contornar este problema, foi criada uma nova etapa, responsável por ajustar os valores obtidos para o plano inicial. Esta etapa soma o tamanho da área recortada acima e à esquerda da área extraída, ao ponto obtido.

Por exemplo, uma imagem de 300 pixels de altura por 300 pixels de largura é recortada para processamento de maneira que a faixa central é selecionada, ou seja, uma nova imagem de apenas 150 pixels é gerada, porém esta faixa tem início na posição 75 da imagem original no eixo Y. Desta maneira, devemos somar 75 ao ponto resultante da extração.

4.2.3.2.1 Extração dos dados do Crânio

Considerando que as técnicas de captura do vídeo tenham sido cumpridas, é possível considerar a afirmação de que o crânio do atleta sempre está no quadrante superior da imagem recebida. Portanto, a área de interesse a ser analisada da imagem refere-se ao primeiro 1/4 da imagem.

Após o recorte da área a ser processada, o sistema deve iniciar a extração dos dados. Para isto, consideramos as seguintes afirmações:

- a) O crânio é sempre o ponto mais alto da imagem.
- b) Partindo da esquerda para a direita e de cima para baixo, vamos encontrar o centro do topo da cabeça do atleta.

c) Partindo do centro, do topo da cabeça, se traçarmos uma reta a 225 graus e outra a 315 graus, basta só localizar o primeiro ponto branco: este ponto irá representar a nuca e o centro da face, especificamente.

Partindo dessas afirmações, o sistema deve percorrer a imagem conforme descrito, e localizar os pontos comentados. Traçando uma reta entre o ponto da nuca e o ponto do centro da face, podemos marcar o meio deste segmento e considerar este ponto como sendo o centro do crânio. Após marcar a imagem, bastará anotar a posição obtida.

4.2.3.2.2 Extração dos dados da Cintura

O tratamento da cintura, vista lateralmente, é feito aplicando-se técnicas semelhantes às vistas anteriormente na extração de dados do crânio; porém, para a cintura, devemos considerar o seguinte cenário:

- a) A cintura está sempre na região centro-inferior da imagem.
- b) A cintura é a base da coluna.
- c) Se considerarmos apenas o segundo terço da imagem, os glúteos irão representar a região mais à direita da imagem.

Com base nestas afirmações, para desenvolver o algoritmo, devemos começar fazendo o recorte da imagem original em 3 pedaços na vertical e selecionar o do meio. Nesta área, devemos localizar o ponto mais extremo à direita, em seguida seguir uma linha reta horizontal e localizar o último ponto à esquerda nesta linha e, então, definir o centro entre eles. Este ponto identifica a cintura.

4.2.3.2.3 Extração dos dados do Joelho

A área de recorte do joelho é representada pelo terceiro quarto da imagem, desde que particionada na vertical. Nesta área o algoritmo deve localizar o ponto mais extremo à esquerda da imagem, que representa a ponta do joelho. Em seguida, o sistema deve localizar o centro entre este ponto e o último ponto à direita.

4.2.3.2.4 Extração dos dados do Pé

O pé está sempre localizado na parte inferior da imagem, desta maneira, basta encontrar o primeiro ponto da direita para a esquerda de baixo para cima e temos a ponto de trás do pé; repetindo o procedimento, apenas invertendo a ordem da esquerda para a direita, teremos o ponto da frente.

4.2.4 Construção de um Modelo Vetorizado

Os pontos extraídos na última fase representam os pontos das partes humanas num plano. A conexão destes pontos representa um desenho vetorizado do atleta capturado.

4.3 ANÁLISE DA CENA

Uma vez obtido o modelo vetorizado do atleta no quadro determinado, o sistema deve executar uma série de verificações matemáticas para garantir que a execução do exercício foi perfeita. Estes cálculos são feitos baseados na distância entre os pontos.

Conforme descrito nos capítulos anteriores, existem alguns parâmetros que garantem a execução do exercício; para os propósitos deste documento, levaremos em consideração alguns deles.

4.3.1 Postura dos Joelhos

Relembrando a regra dos joelhos: “os joelhos não devem ultrapassar a linha do pé”, ou seja, a ponta do joelho não deve ultrapassar a ponta do pé durante todos os momentos da execução. Sendo assim, o sistema deve verificar a posição do joelho no eixo X e da ponta dos pés também no eixo X. A subtração dessas posições não deve nunca retornar um resultado diferente de zero. Caso isto ocorra, significa que a execução não foi realizada com perfeição.

$$\text{joelho}[x] - \text{pe}[x] > 0$$

O valor obtido após o cálculo anterior não representa um grau de perfeição da execução: sempre que o resultado desta subtração for igual a zero, caracterizará uma execução perfeita, e no caso de qualquer resultado diferente de 0, devemos utilizar a distância do joelho para a cintura como parâmetro máximo.

4.3.2 Postura da Cintura

De acordo com os parâmetros definidos para uma boa execução, o atleta não deve fazer um movimento em que a cintura desça abaixo da linha do joelho. O grau de execução para esta regra é gerado através da subtração da posição no eixo Y entre os pontos da cintura e do joelho. Qualquer valor negativo caracteriza uma falha na execução; por outro lado, o sistema deve levar em consideração o caso de o valor das distâncias também ficar sempre muito elevado, o que caracteriza uma execução curta. O parâmetro máximo é sempre a diferença dos pontos entre a cintura para o pé em relação ao eixo Y.

4.3.3 Postura da Coluna

Uma coluna ereta também é um dos parâmetros utilizados. Conforme visto anteriormente, existe um conceito chamado de movimento rígido que é caracterizado por um objeto onde as distâncias entre seus pontos nunca são afetadas, mesmo que estes mudem de posição. Basicamente, será aplicado este conceito para verificar se a coluna permanece ereta e rígida. O sistema deve verificar sempre a distância entre os pontos da cintura e crânio e, caso esta distância sofra alterações, uma falha na execução será caracterizada.

Pode-se utilizar como parâmetro mínimo para esta avaliação uma variação igual a zero como sendo perfeito, e uma variação igual à distância original da coluna como índice máximo.

5 CONCLUSÕES E TRABALHOS FUTUROS

Os objetivos e esforços deste trabalho foram direcionados para a construção de um sistema capaz de realizar a análise do movimento de agachamento com o uso de equipamentos domésticos de maneira simples e precisa.

No que se refere à captura dos movimentos sem uso de marcadores, identificou-se que o emprego de um sistema especialista na extração das características humanas gerou uma limitação durante a etapa de obtenção do vídeo. Esta limitação ocorreu devido à utilização de um mecanismo orientado por um conjunto de regras fixas em código. Percebeu-se que algumas variações não previstas, como a direção do atleta enquanto executa o movimento, pode gerar, em alguns casos, resultados imprecisos. A adição de novas regras prevendo novas possibilidades mostrou-se eficaz na redução da variação dos resultados, porém, diante a milhares de variações de movimento do corpo humano, a enumeração de todas estas em caráter absoluto se mostrou inviável. No entanto, o escopo deste trabalho é claro ao afirmar que algumas limitações são previstas acerca da execução do movimento durante a sua captura; desta maneira, é correto afirmar que os resultados alcançados estão aderentes às premissas deste projeto, e que as considerações realizadas devem ser consideradas como base para futuros trabalhos.

Conforme visto anteriormente, o destacamento da silhueta humana do plano de fundo é crítico para a posterior identificação das características humanas e este problema foi resolvido com a aplicação de dois filtros combinados. Ajustes de acordo com a intensidade de luz, quando realizados apropriadamente, tornaram a silhueta humana mais precisa. Tal parametrização foi realizada manualmente e um valor padrão que atendesse aos casos previstos no escopo deste projeto foi fixado; contudo, identificou-se que a elaboração de um algoritmo para a quantificação da luminosidade pode aumentar a precisão dos dados obtidos.

Durante a conceituação do projeto, ficou definido que o sistema deveria ser alimentado com um vídeo capturado com apenas uma câmera de uso caseiro. Isso garantiu que o sistema não necessitasse de um alto investimento por conta da aquisição de uma série de câmeras digitais e/ou câmeras especiais. Este requisito não funcional foi contemplado com o auxílio de uma biblioteca, Java Media Framework, que permitiu que fossem processados os vídeos submetidos desde que estes vídeos atendessem a algumas especificações quanto ao formato do vídeo enviado; no entanto, estas limitações não implicaram de maneira tão significativa que fosse suficiente para considerarmos que esta meta não foi alcançada.

Uma vez que os dados de entrada sejam precisos, a análise do movimento proporcionou informações bastante conclusivas sobre a qualidade da execução dos movimentos. De posse destes dados, a análise se mostrou precisa em todos os seus aspectos.

O gerenciamento do conteúdo do usuário do sistema se deu com a criação de uma interface web. As interfaces se mostraram bem independentes do restante do sistema e pode-se considerá-las como um módulo complementar do sistema.

Observa-se com isso que todas as premissas deste projeto foram atendidas em sua completude; no entanto, conforme comentado anteriormente, algumas características do sistema foram determinantes para uma redução na flexibilidade do sistema e levantaram algumas questões quanto a trabalhos futuros. Ficou claro que a extração das características humanas foi prejudicada pela escolha de uma solução orientada à especialização de código e que o estudo de outras formas para se resolver este problema podem-se mostrar de grande valia para o aprimoramento do sistema.

Todo o código gerado durante a construção deste trabalho está aberto e disponibilizado no *Google Code* através do endereço <http://code.google.com/p/jmot/source/browse/>. O sistema não foi finalizado, porém encontra-se em estágio avançado. A integração dos processadores de vídeo, processadores de imagens e o sistema web não foram feitas. Porém todos os resultados citados foram obtidos através de testes unitários.

REFERÊNCIAS BIBLIOGRÁFICAS

CAPPELARO, Ap. Esmeraldo; MENDONÇA, C. Fernando; TEIXEIRA, Jeyson. *Você sabe o que é uma câmera digital ?* Disponível em: <<http://educar.sc.usp.br/licenciatura/trabalhos/camera.htm>> Acesso em 01 Agosto. 2009.

DALY, Tim. *Guia Básico de Fotografia Digital*, Lisboa: Estampa, 2002, p. 19.

DAVIS, Larry. *Machine Vision Systems*, Disponível em: <<http://www.umiacs.umd.edu/~lsd/426/426.html>>. Acesso em: 05 abr. 2009.

DAVIS, Larry. *Multi-perspective analysis of human action*. Disponível em: <<http://www.cs.umd.edu/users/thanarat/Research.html>>. Acesso em: 05 abr. 2009.

DELAVIER, Frédéric. *Guia dos Movimentos de Musculação: Abordagem Anatômica*, São Paulo: Manole, 2000, v. 2, p. 4, p. 78-79.

FILHO, Ôge Marques; NETO, Hugo Vieira. *Processamento Digital de Imagens*. Rio de Janeiro: Brasport, 1999, p. 23.

FISHER, Robert; PERKINS, Simon. *Hypermedia Image Processing Reference*. Disponível em: <<http://homepages.inf.ed.ac.uk/rbf/HIPR2/pntops.htm>> Acesso em: 05 abr. 2009.

GAMA, Paola; SENDRA, Fernanda. *A fotografia sequencial de Eadweard Muybridge e o cinema da animação*. Disponível em: <http://www.dad.puc-rio.br/dad07/arquivos_downloads/32.pdf>, Acesso em: 05 abr. 2009.

GENTIL, Paulo. *Agachamento e joelho*. Disponível em: <http://www.educacaofisica.org/joomla/index.php?option=com_content&task=view&id=322&Itemid=2>, Acesso em: 19 mar. 2009.

GOMES, Jonas; VELHO, Luiz. *Computação Gráfica: Imagem*, SBM e IMPA 1994, São Paulo, p. 2, 3.

GONZALES, Rafael C.; WOODS, Richard. *Processamento de Imagens Digitais*, São Paulo: Blucher, 1992, p.4, 26.

NMeM, Science & Society Thomas B; Granum, Erik. *Eadweard Muybridge*. Disponível em: <<http://www.wildfilmhistory.org/person/180/photo/498/Eadweard+Muybridges+time-lapse+photographs+of+a+man+riding+a+galloping+horse.html>>, Acesso em: 03 abr. 2009.

MOESLUND Thomas B; Granum, Erik. *A Survey of Computer Vision-Based Human Motion Capture*. Disponível em: <<http://eprints.kfupm.edu.sa/21400>>, Acesso em: 18 abr. 2009.

OLIVEIRA, Elke. *Musculação: Ponto de vista – Agachamento*, 2002. Disponível em: <http://www.saudeemovimento.com.br/conteudos/conteudo_frame.asp?cod_noticia=843>, Acesso em: 19 mar. 2009.

OLIVEIRA, Elke. *Agachamento e coluna*, 2008. Disponível em: <http://educacaofisica.org/joomla/index.php?option=com_content&task=view&id=323&Itemid=2>, Acesso em: 19 mar. 2009.

PERS, Janez; BON, Marta; KOVACIC, Stanislav; CILIBA, Marko; DEZMAN, Branko. *Observation and analysis of large-scale human*, *Human Movement Science*, p. 4-8.

FISHER, Robert; PERKINS, Simon; WALKER, Ashley; WOLFART, Erik. *Image Processing Learning Resources*. Disponível em: <<http://homepages.inf.ed.ac.uk/rbf/HIPR2/histogram.htm>>, Acesso em 15 mai. 2009.

REMONDINO, Fabio. *Markerless Motion Capture from Single or Multi-Camera Video Sequence*. Zermatt – Switzerland: CAPTECH, 2004, p.1-5.

RIPPETOE, Mark; LON, Kilgore. *Starting Strength – Basic Barbell Training*. 2º Ed. USA – Buchanan: The Aasgaard Company, 2007. 31 p.

SIDNEY, Dwight. *Squat today, grow every day: do this king of all exercises and sprout muscle even in your upper body*, 2003. Disponível em: <http://findarticles.com/p/articles/mi_m1608/is_4_19/ai_104682197/>, Acesso em: 10 mar. 2009.

TAVARES, João Manuel R. S. *Análise do Movimento não rígido em Visão por Computador*. São Paulo: 2004, p. 2-3.

TRACY, V.; WILSON, K.; Nice, G. Gurevich. *HowStuffWorks - Como funcionam as câmeras digitais*. Disponível em: <<http://eletronicos.hsw.uol.com.br/cameras-digitais.htm>>, Acesso em: 05 abr. 2009.