

UNIVERSIDADE FEDERAL DO AMAZONAS  
FACULDADE DE TECNOLOGIA  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

**MONITORAMENTO DE IDOSOS EM  
AMBIENTES INTERNOS POR MEIO DA  
ANÁLISE SINTÁTICA DE DADOS VISUAIS**

José do Patrocínio dos Santos Silva

MANAUS-AM  
2017

UNIVERSIDADE FEDERAL DO AMAZONAS  
FACULDADE DE TECNOLOGIA  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

# **MONITORAMENTO DE IDOSOS EM AMBIENTES INTERNOS POR MEIO DA ANÁLISE SINTÁTICA DE DADOS VISUAIS**

José do Patrocínio dos Santos Silva

Dissertação apresentada ao programa de pós-graduação em Engenharia Elétrica da Universidade Federal do Amazonas, área de concentração Sistemas Inteligentes e Engenharia de Software, parte dos requisitos para obtenção do título de mestre em Engenharia Elétrica.

Orientador: Prof. Dr. José Luiz de Souza Pio

MANAUS-AM  
2017

FEDERAL UNIVERSITY OF AMAZONAS  
FACULTY OF TECHNOLOGY  
GRADUATE PROGRAM IN ELECTRICAL ENGINEERING

**MONITORING OF ELDERLY IN INDOORS  
ENVIRONMENTS THROUGH VISUAL  
DATA SYNTATIC ANALYSIS**

José do Patrocínio dos Santos Silva

Dissertation presented to the graduate program in Electrical Engineering of Federal University of Amazonas, concentration area Intelligent Systems and Software Engineering, part of the requirements for obtaining the Master's degree in Electrical Engineering.

Advisor: Dr. José Luiz de Souza Pio

MANAUS-AM  
2017

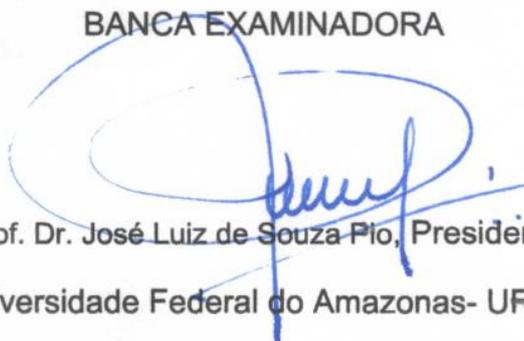
JOSÉ DO PATROCÍNIO DOS SANTOS SILVA

MONITORAMENTO DE IDOSOS EM AMBIENTES INTERNOS POR MEIO DA  
ANÁLISE SINTÁTICA DE DADOS VISUAIS

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Amazonas, como requisito parcial para obtenção do título de Mestre em Engenharia Elétrica na área de concentração Sistemas Inteligentes e Engenharia de Software.

Aprovado em 05 de abril de 2017.

BANCA EXAMINADORA



Prof. Dr. José Luiz de Souza Pío, Presidente

Universidade Federal do Amazonas- UFAM



Prof. Dr. Waldir Sabino da Silva Júnior, Membro

Universidade Federal do Amazonas- UFAM



Profª. Dra. Eulanda Miranda dos Santos, Membro

Universidade Federal do Amazonas- UFAM

## Ficha Catalográfica

Ficha catalográfica elaborada automaticamente de acordo com os dados fornecidos pelo(a) autor(a).

S237m Santos Silva, José do Patrocínio dos  
Monitoramento de idosos em ambientes internos por meio da  
análise sintática de dados visuais / José do Patrocínio dos Santos  
Silva. 2017  
85 f.: il. color; 31 cm.

Orientador: José Luiz de Souza Pio  
Dissertação (Mestrado em Engenharia Elétrica) - Universidade  
Federal do Amazonas.

1. Monitoramento de idosos. 2. gramática regular. 3. análise  
sintática. 4. queda de idosos. I. Pio, José Luiz de Souza II.  
Universidade Federal do Amazonas III. Título

## AGRADECIMENTOS

Aos meus familiares por todo amor dedicação e incentivo a minha educação.

Ao meu orientador Prof. José Luiz de Souza Pio por toda paciência, dedicação e amizade.

A todos os meus Ex-Professores da UFAM e demais instituições em que estudei: UFCG, Colégio Estadual José Bronzeado Sobrinho e Escola Municipal Antônio Carneiro Filho ambas em Remígio-PB.

A todos os meus ex-colegas de turma do PPGEE da UFAM e também do DEE da UFCG, pelo conhecimento compartilhado e bons momentos vivenciados.

A todos os funcionários do PPGEE e ICOMP da UFAM e também da UFCG, os quais eu sempre pude contar quando necessitei e pela amizade adquirida.

A todos(a) amigos(a) da minha cidade natal Remígio-PB, de Campina Grande-PB e de Manaus-AM, por todo companheirismo.

A meus colegas de trabalho da Samsung e meus superiores, pelo incentivo e disposição em atender minhas solicitações sempre que necessitei.

A todos que torceram por mim e contribuíram para este momento.

A CAPES pela concessão da bolsa.

## RESUMO

Nos dias atuais, os sistemas para monitoramento de atividades humanas são utilizados para detectar situações de risco ou atividades que possam comprometer a integridade física ou a saúde das pessoas. Esses sistemas favorecem a atuação ou intervenção mais rápida e eficiente de profissionais responsáveis pela guarda e monitoramento de pessoas idosas, crianças ou enfermos a partir da detecção e reconhecimento de ações ou atividades de risco. Este trabalho objetiva realizar o monitoramento e reconhecimento de queda de pessoas idosas em ambiente interno provido por um sistema de câmeras. A grande incidência de queda de idosos tem sido considerada um dos principais problemas de saúde pública e afeta diretamente a economia devido aos custos consideráveis de internação, tratamento e assistenciais. A abordagem desenvolvida utiliza descrições sintáticas de primitivas extraídas e caracterizadas a partir de movimentos de idosos. Esta metodologia foi aplicada em diversas situações cotidianas de pessoas idosas que eventualmente sofrem uma queda em ambientes internos. Os resultados obtidos demonstram-se satisfatórios e apontam para a viabilidade e integração da metodologia aos sistemas de monitoramento de atividades de pessoas idosas.

Palavras Chaves: Monitoramento de idosos, gramática regular, análise sintática, queda de idosos.

## **ABSTRACT**

Systems for monitoring human activities are used to detect risk situation that may compromise one's physical integrity or health. These systems support more efficient actions from people responsible for guarding and monitoring the elderly, children, or ill through identification of dangerous activities. This work aims to perform fall recognition of elderly people indoors properly monitored by a camera system. The developed approach uses syntactic descriptions of primitives extracted and characterized from the movements of the elderly. The recognition is given by finite-state automata governed by pre-established grammatical rules. This methodology was applied in several daily situations of elderly people who eventually suffer a fall indoors. The results obtained prove to be satisfactory and point to the feasibility and integration of the methodology in the monitoring systems of elderly activities.

**Keywords:** Elderly activities monitoring, grammar, syntactic analysis, fall recognition.

## LISTA DE ILUSTRAÇÕES

Figura 1 - Exemplo de um ambiente interno monitorado por câmeras .....	15
Figura 2 - Diferentes fases de uma queda .....	15
Figura 3 - Movimento de uma pessoa andando decomposto em primitivas visuais .....	16
Figura 4 - Transição entre dois estados de um autômato .....	24
Figura 5 – Monitoramento de uma pessoa e sua respectiva silhueta.....	25
Figura 6 - Aproximação da silhueta humana por uma elipse .....	26
Figura 7 - Elipse aproximada para diferentes posturas humana.....	27
Figura 8 – Exemplo de segmentação e construção da gramática .....	28
Figura 9 - Fluxograma do processo de treinamento .....	30
Figura 10 - Fluxograma do processo de reconhecimento.....	31
Figura 11 - Resultado da subtração entre <i>Frames</i> .....	33
Figura 12 - Imagem binária da silhueta ilustrada na Figura 11 (c) .....	34
Figura 13 - Fechamento da imagem da silhueta da Figura 12.....	35
Figura 14 - Silhueta da forma humana gera uma região que é aproximada por uma elipse.....	36
Figura 15 - Semi-eixo $b$ , eixo maior $a$ , centro $(x, y)$ e inclinação $\theta$ da elipse.....	36
Figura 16 - Elipse paralela aos eixos: $Y$ e $X$ .....	38
Figura 17 - Elipse aproximada da silhueta de uma pessoa andando (a) e caída (b) .....	38
Figura 18 - Zonas de posição: (a) andando, (b) transição e (c) caído .....	39
Figura 19 - <i>Frames</i> da queda de um idoso que estava deambulando.....	46
Figura 20 - Silhuetas das subtrações dos <i>frames</i> da Figura 19 e suas respectivas inclinações .....	47
Figura 21 - Posições/movimentos versus os caracteres da aplicação 01.....	49
Figura 22- Distribuição das silhuetas por zonas de posição durante o treinamento.....	50
Figura 23 - Autômato $A = (Q, \Sigma, \delta, S, T)$ .....	51
Figura 24 - <i>Frames</i> da queda de um idoso derrubado pelo seu cão .....	52
Figura 25 - Silhuetas das subtrações dos <i>frames</i> da Figura 24 e suas respectivas inclinações .....	53
Figura 26 - Distribuição das silhuetas por zonas durante a aplicação 01 .....	54

Figura 27 - Posições/movimentos versus caracteres obtidos da aplicação 01 .....	55
Figura 28 - <i>Frames</i> da queda de um idoso durante prática de exercício .....	56
Figura 29 - Silhuetas das subtrações dos <i>frames</i> da Figura 28 e suas respectivas inclinações .....	57
Figura 30 - Distribuição dos Silhuetas por zonas durante a aplicação 02 .....	58
Figura 31 - Posições/movimentos obtidos da aplicação 02.....	59
Figura 32 - <i>Frames</i> da queda de um idoso após tropeçar em um carrinho de limpeza .....	60
Figura 33 - Silhuetas das subtrações dos <i>frames</i> da Figura 32 e suas respectivas inclinações .....	61
Figura 34 - Distribuição das Silhuetas por zonas durante a aplicação 03 .....	62
Figura 35 - Posições/movimentos obtidos da aplicação 03.....	63
Figura 36 - Idoso cai ao levantar da cadeira .....	64
Figura 37 – Silhuetas obtidas da subtração dos <i>frames</i> da Figura 36.....	65
Figura 38 - Distribuição das silhuetas por zonas durante a aplicação 04 .....	66
Figura 39 - Posições/movimentos obtidos da aplicação 05.....	67
Figura 40 - Idoso cai ao descer escadaria do ICOMP .....	68
Figura 41 - Silhuetas obtidas da subtração dos <i>frames</i> da Figura 40 .....	69
Figura 42 - Distribuição das silhuetas por zonas durante a aplicação 05 .....	70
Figura 43 - Posições/movimentos obtidos da aplicação 05.....	71
Figura 44 - <i>Frames</i> da aplicação realizada por Foroughi <i>et al.</i> (2008) .....	72
Figura 45 - Silhuetas e suas respectivas inclinações dos <i>frames</i> da Figura 44.....	73
Figura 46 - Distribuição das silhuetas por zonas durante a aplicação 06.....	74
Figura 47 - Posições/movimentos obtidos da aplicação 06.....	74

## LISTA DE TABELAS

Tabela 1 – Procedimento experimental.....	45
Tabela 2 - Parâmetros obtidos da fase de treinamento.....	48
Tabela 3 - Caracteres obtidos do treinamento.....	49
Tabela 4 - Palavras que representam as Posições e Movimentos.....	50
Tabela 5 - Caracteres obtidos da aplicação 01.....	54
Tabela 6 - Matriz de confusão referente aos resultados da aplicação 01.....	55
Tabela 7 - Caracteres obtidos da aplicação 02.....	58
Tabela 8 - Matriz de confusão referente aos resultados da aplicação 02.....	59
Tabela 9 - Caracteres obtidos da aplicação 03.....	62
Tabela 10 - Matriz de confusão referente aos resultados da aplicação 03.....	63
Tabela 11 - Caracteres obtidos da aplicação 04.....	66
Tabela 12 - Matriz de confusão referente aos resultados da aplicação 04.....	67
Tabela 13 - Caracteres obtidos da aplicação 05.....	70
Tabela 14 – Matriz de confusão referente aos resultados da aplicação 05.....	71
Tabela 15 - Caracteres obtidos da aplicação 06.....	73
Tabela 16 – Matriz de confusão referente aos resultados da aplicação 06.....	75
Tabela 17 - Resultado da comparação entre os métodos.....	75
Tabela 18 - Matriz de confusão referente ao total de silhuetas.....	76
Tabela 21 - Métricas para avaliação do método e valores obtidos.....	77

# SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO.....</b>	<b>14</b>
1.1	O PROBLEMA E SUA CONTEXTUALIZAÇÃO .....	14
1.2	HIPÓTESE.....	16
1.3	MOTIVAÇÃO.....	17
1.4	JUSTIFICATIVA.....	18
1.5	CONTRIBUIÇÕES .....	19
1.6	OBJETIVOS.....	20
1.6.1	Geral.....	20
1.6.2	Específico.....	20
1.7	ORGANIZAÇÃO DO TRABALHO .....	20
<b>2</b>	<b>REVISÃO BIBLIOGRÁFICA .....</b>	<b>21</b>
2.1	FUNDAMENTAÇÃO TEÓRICA .....	21
2.1.1	Linguagens formais e Análise Sintática .....	21
2.1.2	Autômatos Finitos Determinísticos .....	23
2.2	TRABALHOS RELACIONADOS .....	25
<b>3</b>	<b>MONITORAMENTO SINTÁTICO DE MOVIMENTOS DE IDOSOS .....</b>	<b>29</b>
3.1	FORMAÇÃO DAS SENTENÇAS GRAMATICAIIS (TREINAMENTO) .....	29
3.2	RECONHECIMENTO. ....	31
3.2.1	1ª Etapa: Aquisição de dados visuais .....	32
3.2.2	2ª Etapa: Formação das palavras .....	32
3.2.3	3ª Etapa: Reconhecimento sintático .....	41
<b>4</b>	<b>APLICAÇÕES EXPERIMENTAIS.....</b>	<b>42</b>
4.1	PROCEDIMENTOS E MÉTODOS.....	42
4.1.1	Métricas.....	42

4.2	PLANEJAMENTO EXPERIMENTAL.....	45
4.2.1	Realização do treinamento e definição da base de regras .....	45
4.2.2	Aplicação 01: Idoso é derrubado pelo seu cão .....	51
4.2.3	Aplicação 02: Idoso sofre queda ao praticar exercício em sua casa.....	55
4.2.4	Aplicação 03: Idoso cai ao tropeçar em objeto em sua casa .....	59
4.2.5	Aplicação 04: Idoso cai ao tentar levantar da cadeira .....	63
4.2.6	Aplicação 05: Idoso cai ao descer escada.....	67
4.2.7	Aplicação 06: Comparação com outros métodos .....	71
4.3	ANÁLISE DOS EXPERIMENTOS.....	75
<b>5</b>	<b>CONCLUSÕES E TRABALHOS FUTUROS.....</b>	<b>79</b>
5.1	CONCLUSÕES.....	79
5.2	SUGESTÕES PARA TRABALHOS FUTUROS.....	80
5.3	CONSIDERAÇÕES FINAIS .....	81
	<b>REFERÊNCIAS.....</b>	<b>82</b>

---

# INTRODUÇÃO

Monitoramento de atividades humanas tem sido um tema de pesquisa recorrente nas áreas científicas de Visão Computacional, Aprendizado de Máquinas e Inteligência Artificial. Isso devido o fato da sociedade necessitar cada vez mais de meios que favoreçam a segurança, saúde e conforto das pessoas (CANDAMO *et al.*, 2010; KE *et al.*, 2013). Este trabalho aborda o reconhecimento das atividades de idosos em ambientes internos por meio de uma metodologia baseada no reconhecimento sintático de informações visuais para a detecção de queda de pessoas idosas.

## 1.1 O PROBLEMA E SUA CONTEXTUALIZAÇÃO

Devido às complicações para saúde, custos assistenciais e sua alta incidência, as quedas sofridas por pessoas idosas são consideradas como um dos principais problemas de saúde pública (LUZ *et al.*, 2013). Guimarães *et al.* (2004) definem queda como um deslocamento não intencional do corpo para um nível inferior à posição inicial, com incapacidade de correção em tempo hábil. Resultam da interação entre diversos fatores que comprometem a estabilidade do indivíduo, tais como: ambientais, fisiológicos, psicossociais, biomédicos, entre outros.

O problema científico abordado neste trabalho objetiva monitorar e reconhecer situações que ocasionem a queda de pessoas idosas em ambientes internos a partir de dados visuais obtidos por câmeras convenientemente dispostas. As câmeras cobrem áreas como sala de estar, escadas, cozinha, entrada do banheiro, quarto, entre outros. A Figura 1 ilustra um ambiente interno típico, monitorado por um sistema de câmeras em circuito fechado.

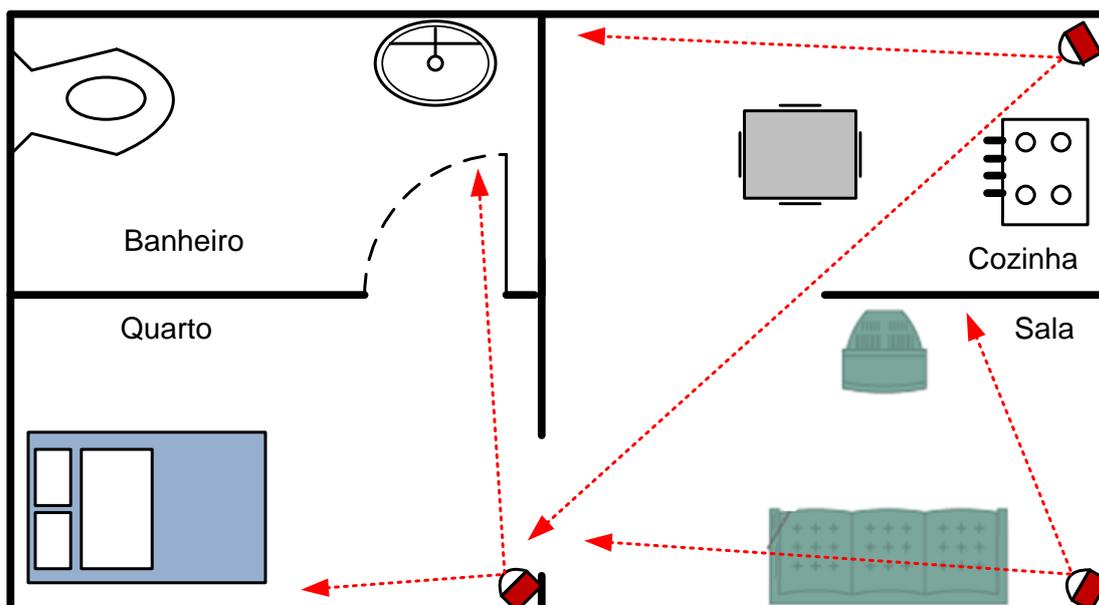


Figura 1 - Exemplo de um ambiente interno monitorado por câmeras.

Fonte: Autor.

A principal dificuldade em detectar a queda de idosos é reconhecer esse movimento entre todas as atividades da vida diária, especialmente sentar-se e agachar-se que são movimentos que têm características semelhantes. Rougier *et al.* (2011) dividem a queda em quatro fases: pré-queda, fase crítica, pós-queda e fase de recuperação. Estas fases estão ilustradas na Figura 2.

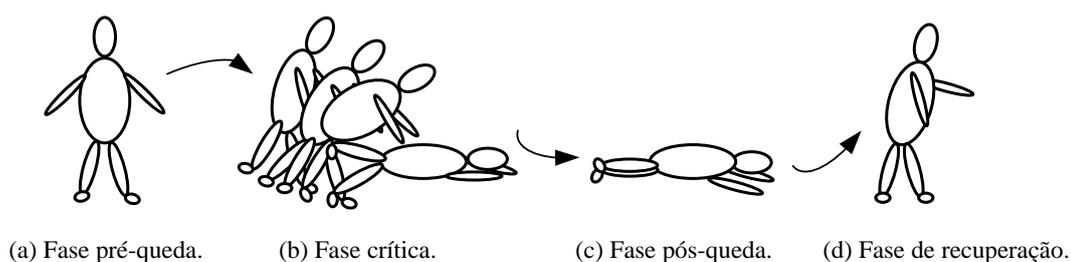


Figura 2 - Diferentes fases de uma queda.

Fonte: adaptado de (ROUGIER *et al.*, 2011).

A fase pré-queda, ilustrada na Figura 2 (a), corresponde a movimentos da vida diária, seguida de movimentos bruscos, ocasionalmente, dirigidos para o chão como sentado ou agachado. Essas atividades não devem gerar o alarme do sistema de monitoramento. A Figura 2 (b) ilustra a fase crítica, que corresponde à queda propriamente dita. Essa fase geralmente é extremamente curta, pode ser detectada pelo movimento do corpo ou pelo choque de impacto com o solo. A fase pós-

queda, ilustrada na Figura 2 (c) é geralmente caracterizada pela imobilidade da pessoa no chão logo após a queda e pode ser detectada por uma posição deitada ou por uma ausência significativa de movimento. Já a Figura 2 (d) ilustra a fase de recuperação que pode eventualmente ocorrer se a pessoa levantar sozinha, ou com a ajuda de outros.

## 1.2 HIPÓTESE

Assume-se como hipótese que o reconhecimento da queda de idosos em ambientes internos, pode ser alcançado por meio da descrição sintática de padrões visuais obtidos a partir de sequências de vídeo.

Considera-se que padrões visuais em uma sequência de imagens associa cada movimento a um símbolo. A Figura 3 ilustra o movimento de uma pessoa andando decomposto em termos de suas primitivas visuais e seus respectivos símbolos.

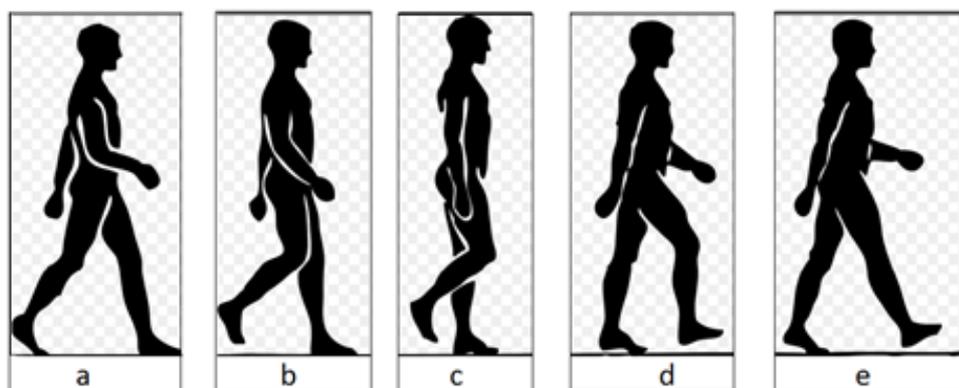


Figura 3 – Movimento de uma pessoa andando decomposto em primitivas visuais.  
Fonte: lenecoaching.com.br.

Por analogia formal com a sintaxe de uma linguagem, pode-se admitir esses padrões como sentenças pertencentes a um determinado alfabeto. Essas sentenças são geradas a partir de extração de características. Então um conjunto de regras denominadas *regras gramaticais* são combinadas para definir um determinado movimento. Logo, a geração de sequências de padrões representa um determinado movimento característico, como a queda e suas fases anteriormente representadas.

Essa abordagem enfatiza a descrição estrutural de padrões visuais para o reconhecimento de queda de idosos. Trata-se de uma abordagem matematicamente consistente por causa da disponibilidade da teoria das linguagens formais e dos autômatos como fundamentação.

Se o movimento da pessoa idosa pode ser decomposta em uma sequência de padrões visuais, então uma gramática ou um autômato regular é capaz de reconhecer o tipo de movimento executado.

### **1.3 MOTIVAÇÃO**

O mundo está enfrentando um crescimento demasiado da população idosa, devido o aumento da expectativa de vida. No Canadá, por exemplo, em 2001 um em cada oito canadenses tinha mais de 65 anos de idade, esta proporção será de um em cada cinco em 2026 (ROUGIER *et al.*, 2011). As pessoas idosas tendem a passar a maior parte de seu tempo em casa, daí existe um interesse na pesquisa de meios para garantir a segurança e saúde de pessoas idosas em casa. As quedas são um dos principais riscos para os idosos que vivem sozinhos em casa, muitas vezes causando danos irreparáveis a sua integridade física. O risco é amplificado se o idoso não pode pedir ajuda.

No Brasil, a prevalência de quedas chega a 27,6%, isto é um fato preocupante, visto que a população de idosos no Brasil tende a crescer, devido o avanço da medicina e da preocupação das pessoas com uma vida mais saudável, que contribui de forma significativa para o aumento da longevidade no Brasil (SIQUEIRA *et al.*, 2011; MICHELETTO, 2011; ROZESTRATEN, 1993; ALVES, 2010). A média de vida da população brasileira era de 45,5 anos, em 1940, passando para 75,5 anos, em 2015. Com o aumento da expectativa de vida, a média de vida do brasileiro, em 2050, será de 81,29 anos, ou seja, aproximadamente 1 em cada 4 brasileiros terá mais de 65 anos (IBGE, 2016; LUZ *et al.*, 2013; MORAES *et al.*, 2012; GANANÇA *et al.*, 2006; CARVALHO *et al.*, 2002).

O tempo de atendimento ao idoso após a ocorrência de uma queda, segundo estudos médicos, é extremamente importante tanto na recuperação deste paciente, como também na tentativa de

evitar mortes, além das sequelas que o paciente pode levar para o resto de sua vida por não ter recebido um atendimento rápido e eficiente (SANTOS e ANDRADE, 2005).

Logo, sistemas de monitoramento favorecem o idoso que não quer abrir mão de sua independência e deseja garantir sua autonomia sem perder a segurança, ao mesmo tempo proporciona tranquilidade a seus familiares.

#### 1.4 JUSTIFICATIVA

Reconhecimento de atividades humanas a partir de monitoramento visual é um tema abordado em muitas pesquisas voltadas para aplicações em sistemas de segurança, vigilância, monitoramento de trânsito, transporte de pessoas, monitoramento de idosos, de crianças, de enfermos, em ambientes fabris, na indústria do entretenimento, em ambientes e casas inteligentes (AMOR *et al.*, 2016; KE *et al.*, 2013; CANDAMO *et al.*, 2010; AGGARWAL *et al.*, 2011; AKANSHA *et al.*, 2014; AMOR *et al.*, 2016; DOLLAR *et al.*, 2012; IVANOV e BOBICK, 1998; KO, 2008; DALDOSS, 2013; OGALE *et al.*, 2015; POPPE, 2007; TURAGA *et al.*, 2008; WEINLAND *et al.*, 2011; XIA *et al.*, 2012; YANG *et al.*, 2014; YAO *et al.*, 2009; IJJINAL *et al.*, 2014; GAGLIO *et al.*, 2015; MODARRES *et al.*, 2013). No entanto, os trabalhos citados acima não tratam especificamente sobre reconhecimento de queda de idosos em ambientes internos, já a abordagem proposta neste trabalho trata deste tema, visto a necessidade do desenvolvimento de meios que favoreçam a população idosa que cresce a cada dia.

Alguns trabalhos tratam especificamente sobre o reconhecimento de quedas, como Rougier *et al.* (2011), Worrakulpanit *et al.* (2014), Dennis *et al.* (2008), Tang *et al.* (2015), Noury *et al.* (2008), Sixsmith *et al.* (2004). Alguns destes trabalhos abordam uma metodologia fundamentada em um coeficiente de movimento brusco, que detecta o movimento brusco durante a queda, no entanto, possivelmente não detecta a queda caso o idoso caia lentamente, ou seja, se ele estiver passando mal e apoiar-se em alguma coisa e for caindo lentamente sem movimentos bruscos, então a queda não será detectada. Também fazem uso de dispositivos, por exemplo, o acelerômetro, no entanto, estes dispositivos encarecem a aplicação, pois fazem uso de baterias que necessitam serem

recarregadas e de possíveis substituições, além de proporcionar desconforto ao usuário que precisa estar com o dispositivo detector fixo ao corpo. Outros dispositivos funcionam após a pessoa pressionar um botão de pressão, mas é inútil se o idoso está imobilizado ou inconsciente após a queda. Logo, meios para detecção automática, como o proposto neste trabalho, são mais eficientes, pois não é necessária qualquer intervenção humana e detecta a queda sendo ela brusca ou não, como pode ser averiguado durante as aplicações experimentais em que é reconhecida a queda de idosos sob diversas circunstâncias, pois o método avalia palavras constituídas por cadeias de caracteres obtidos a partir de padrões visuais que representam o movimento do idoso.

## 1.5 CONTRIBUIÇÕES

A principal contribuição deste trabalho é prover uma abordagem que consiste na *Análise Sintática de Dados Visuais* para tratar o problema de reconhecimento de atividades, baseado na descrição sintática por meio de gramática regular. As gramáticas são construídas a partir da decomposição do movimento em um conjunto de primitivas, facilmente manipuláveis e configuráveis para a formação de uma base de reconhecimento de movimentos, atividades e ações humanas, deste modo à queda pode ser reconhecida independente da rapidez, do modo ou posição que o idoso venha a cair, que torna-se uma vantagem em relação a outros trabalhos baseados em coeficientes de movimento brusco. O método é de baixo custo de implementação ao contrário de outros métodos que utilizam dispositivos para reconhecimento da queda.

Por outro lado, o trabalho também favorece integrar o método proposto aos sistemas de monitoramento de idosos já implantados em ambientes internos em que os dados visuais podem ser obtidos a partir das câmeras já instaladas no ambiente, ajudando o profissional responsável pelo monitoramento do idoso ao avisá-lo de situações de risco, em particular das quedas, livrando-o da constante observação do ambiente monitorado.

## 1.6 OBJETIVOS

Para uma melhor organização do texto, os objetivos serão divididos em geral e específicos.

### 1.6.1 Geral

Criar um método, em que a partir de *Caracteres* obtidos de *Padrões Visuais* extraídos do monitoramento de idosos, seja possível identificar a queda de pessoas idosas em um ambiente interno como uma residência, asilo, hospitais ou clínicas, favorecendo o monitoramento e a melhoria no atendimento a pessoas idosas em situações de risco.

### 1.6.2 Específico

- Extrair padrões visuais dos movimentos para obter as primitivas;
- Compor os caracteres a partir das primitivas extraídas;
- Estruturar o modelo em gramáticas formais;
- Reconhecer os movimentos a partir das regras gramaticais estabelecidas.

## 1.7 ORGANIZAÇÃO DO TRABALHO

Esta dissertação de mestrado está dividida em cinco capítulos. O Capítulo 1 trata da introdução onde são apresentados o problema e o objetivo do trabalho. Capítulo 2 aborda a fundamentação teórica e os trabalhos relacionados. O Capítulo 3 trata da abordagem proposta, detalhando o procedimento para extração de primitivas, construção das gramáticas e do método da aproximação da silhueta humana a uma elipse. O Capítulo 4 apresenta as aplicações do método e os respectivos resultados por meio de figuras, gráficos e tabelas. Finalmente em Conclusões e Trabalhos futuros são relatados os pontos relevantes deste trabalho, as limitações, dificuldades obtidas e também sugestões para trabalhos futuros.

---

## REVISÃO BIBLIOGRÁFICA

Este capítulo trata da fundamentação teórica abordada nesta pesquisa e também apresenta o estado da arte das pesquisas sobre reconhecimento de ações humanas e queda de idosos em ambientes internos a partir de dados visuais.

### 2.1 FUNDAMENTAÇÃO TEÓRICA

#### 2.1.1 Linguagens formais e Análise Sintática

Um determinado idioma ou linguagem pode ser definido por cadeias de caracteres compostas de conjuntos finitos de símbolos. A gramática é a ferramenta que permite especificar quais cadeias de caracteres pertencem a uma determinada linguagem (VIEIRA, 2006).

Uma linguagem formal tem sintaxe bem definida, de forma que, dada uma sentença, seja sempre possível saber se ela pertence ou não à linguagem. A linguagem também tem semântica precisa, de modo que não contenha ambiguidades ou sentenças sem significado (VIEIRA, 2006).

Toda linguagem está associada a um alfabeto. Um alfabeto  $\Sigma$  é um conjunto de símbolos, finito e não vazio. Uma palavra ou uma cadeia de símbolos, ou ainda cadeia de caracteres, é uma sequência finita de símbolos (PIO *et al.*, 2006). A sintaxe das linguagens é expressa na forma de uma gramática.

Segundo Vieira (2006), gramáticas são o formalismo projetado para a definição da linguagem. Um formalismo que mostra como gerar as palavras de uma linguagem. O elemento fundamental da gramática é a regra. Uma regra é um par ordenado  $(u, v)$ , em geral escrito da forma  $u \rightarrow v$ , em que  $u$  e  $v$  são palavras de dois alfabetos disjuntos, um com símbolos denominados de variáveis, ou não terminais, e outro com símbolos denominados de terminais. As variáveis são símbolos auxiliares para a geração das palavras da linguagem, enquanto o conjunto de terminais nada mais é do que o

alfabeto da linguagem definida, usando letras maiúsculas para representar as variáveis e minúsculas para os terminais. Formalmente uma gramática é definida como uma quádrupla  $(V, \Sigma, R, P)$ , em que:

- a)  $V$  é um conjunto finito de elementos denominados de variáveis;
- b)  $\Sigma$  é um alfabeto, tal que  $V \cap \Sigma = \emptyset$ ;
- c)  $R \subseteq (V \cup \Sigma)^+ \times (V \cup \Sigma)^*$  é um conjunto finito de pares ordenados chamados de regras;
- d)  $P \in V$  é uma variável conhecida como variável de partida.

Na análise sintática de padrões, o alfabeto  $\Sigma$  é o conjunto de primitivas dos padrões. Em algumas aplicações, a escolha do conjunto de primitivas é muito difícil de ser realizada, residindo aí a maior complexidade dos problemas em reconhecimento sintático. O conjunto de símbolos não terminais  $V$ , ou variáveis, representa a criação ou ocorrência de um padrão específico consistindo exclusivamente de símbolos terminais. O conjunto de produções  $R$ , ou regras de produções ou regras de reescrita, viabiliza a substituição e derivação de símbolos da gramática.

Uma gramática pode ser empregada de dois modos:

1. Generativo, onde a gramática é usada para criar sequências de símbolos terminais (palavras) usando  $R$ , deste modo uma sentença na linguagem da gramática é gerada;
2. Analítico, em que dada uma sentença, junto com a especificação da gramática  $G$ , deseja-se determinar, se a sentença foi gerada por  $G$ .

A linguagem gerada pela gramática  $G$  é o conjunto das formas sentenciais finais (contendo apenas símbolos terminais) deriváveis a partir da variável de partida  $P$ . Estas formas sentenciais finais são chamadas sentenças da linguagem gerada pela gramática, ou de maneira mais sucinta  $L(G)$ .

Esta pesquisa considera que as ações humanas podem ser associadas à sequência de elementos característicos, o reconhecimento destas ações implica em uma análise de conteúdo e contexto destes elementos. Considerando o monitoramento visual, os dados de vídeo contêm uma grande quantidade de informações de características importantes para inferir o comportamento humano, por exemplo, a localização, postura e movimento de uma pessoa. Considerando que estas

características podem representar primitivas, tal problema é passível de ser modelado sintaticamente de acordo com a teoria da gramática formal (PIO, J. L. S. *et al.*, 2006).

### 2.1.2 Autômatos Finitos Determinísticos

Os autômatos finitos constituem um modelo útil para sistemas que a cada momento podem estar em algum estado, de um número finito de estados, que são elementos que definem as características atuais, além de memorizar parte da história do sistema. Conforme o sistema recebe entradas, faz a transição de um estado para outro, até o momento em que deixa de receber entradas. Neste momento, caso o autômato encontre-se em um estado que pertence ao conjunto de estados finais, considera-se que a entrada foi aceita (LINDEN *et al.*, 2009; VIEIRA, 2006).

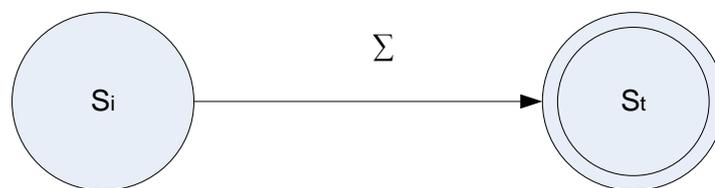
Formalmente, um *Autômato Finito Determinístico-AFD*, pode ser descrito por cinco características:

1.  $Q$  é um conjunto finito não vazio de estados do autômato;
2.  $\Sigma$  é um conjunto de símbolos, denominado alfabeto de entrada do autômato;
3.  $\delta$  é a função de transição de estados do autômato e seu papel é o de indicar as transições possíveis em cada configuração do autômato. Esta função fornece para cada par *estado e símbolo de entrada* um novo estado para onde o autômato deverá mover-se, em que  $\delta: Q \times \Sigma \rightarrow Q$ ;
4.  $S$  é denominado estado inicial do autômato finito. É o estado para o qual o reconhecedor deve ser levado antes de iniciar suas atividades, em que  $S \in Q$ ;
5.  $T$  é um subconjunto do conjunto  $Q$  dos estados do autômato, e contém todos os estados de aceitação ou estados finais do autômato finito. Estes estados são aqueles em que o autômato deve terminar o reconhecimento das cadeias de entrada que pertencem à linguagem que o autômato define. Nenhuma outra cadeia deve ser capaz de levar o autômato a qualquer destes estados, em que  $T \subseteq Q$ .

Assim, pode-se representar totalmente um autômato finito através de uma quintupla  $A = \langle Q, \Sigma, \delta, S, T \rangle$ .

As transições são representadas por triplas  $\{S_i, \Sigma, S_t\}$ , onde  $S_i$  é um estado inicial da transição,  $\Sigma$  é o conjunto de símbolos do alfabeto (caracteres) que disparam esta transição quando o estado corrente é  $S_i$  e  $S_t$  será o novo estado corrente do autômato após a transição. Se para cada elemento de  $\Sigma$  existe apenas uma transição e se nenhuma transição pode ser feita sobre a entrada vazia, então o autômato é dito determinístico (*AFD*). Caso contrário ele é dito Não-Determinístico (*AFND*).

Autômatos são usualmente representados na forma de um grafo dirigido, onde estados são representados por círculos, sendo que estados finais são representados por círculos duplos, e as transições por arestas rotuladas com os símbolos que disparam a transição entre os dois estados conectados, como ilustra a Figura 4.



**Figura 4 - Transição  $\{S_i, \Sigma, S_t\}$ . O círculo do estado  $S_f$  é duplo pois ele pertence ao conjunto  $T$  de estados finais.**

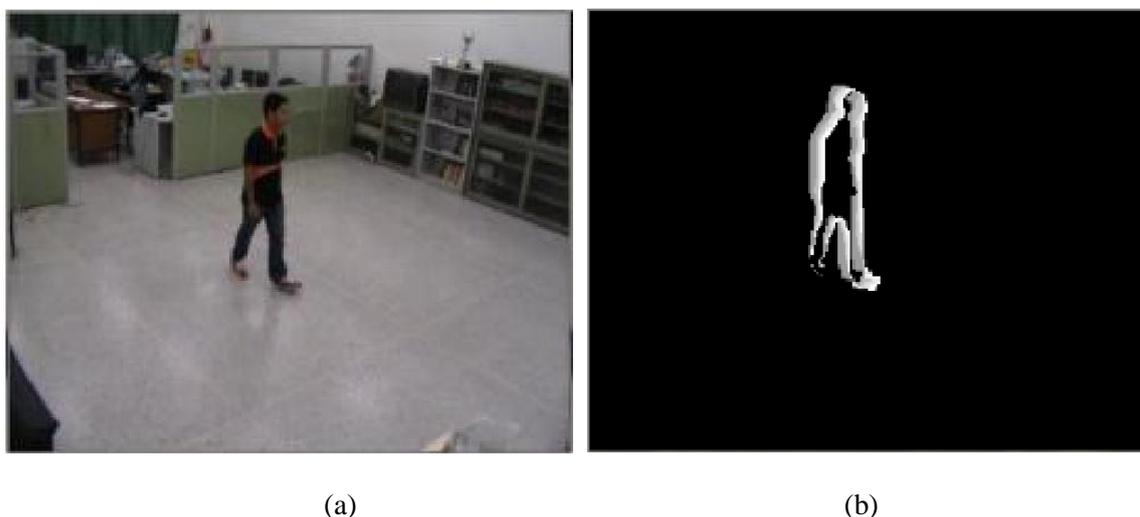
**Fonte: (LINDEN *et al.*, 2009).**

Os Autômatos Finitos Não-Determinísticos diferem dos autômatos finitos determinísticos quanto á regras de transição entre estados. Dada uma combinação de um estado atual e um símbolo de entrada, pode não haver estados especificados para os quais o estado atual deve conduzir o processamento, bem como pode haver vários estados resultantes da leitura do símbolo. Portanto, para uma função de transição definida em  $Qx\Sigma$ , o seu valor não deve ser um elemento de  $Q$  (como acontece com os autômatos determinísticos), mas um subconjunto de  $Q$  (incluindo o conjunto vazio). Ou seja, o processamento de  $\delta(q, a)$  leva á um conjunto de estados em que a máquina pode legalmente se encontrar após estar em um estado  $q$  lendo um símbolo de entrada  $a$ . Embora um *AFND* seja somente um acréscimo ao *AFD*, na verdade não aumenta seu poder computacional, sendo assim, para cada *AFND* é possível construir um *AFD* equivalente que realiza o mesmo processamento. O contrário também é verdadeiro (LINDEN *et al.*, 2009; VIEIRA, 2006).

## 2.2 TRABALHOS RELACIONADOS

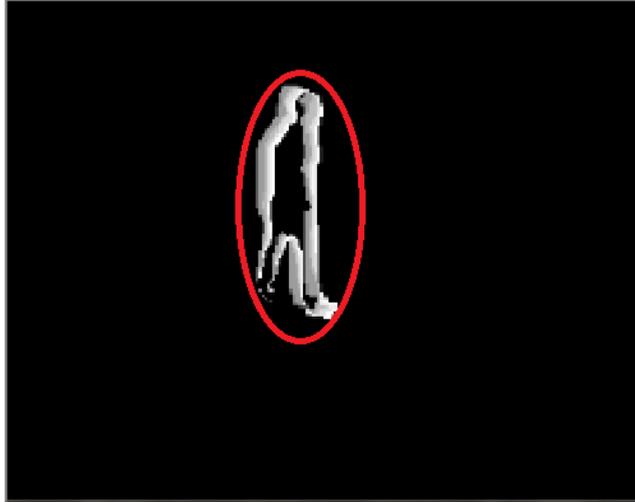
O reconhecimento de queda em ambientes internos é um desafio para a comunidade científica, soluções para este tipo de problema têm sido muito difíceis de serem obtidas, e diversos métodos buscam tratar tal problema.

O método mais usual para o reconhecimento automáticos da queda se dá a partir da detecção da variação brusca de movimentos, como proposto por Rougier *et al.* (2011) e Worrakulpanit *et al.* (2014) que propuseram um método em que a queda é detectada por meio de um rápido movimento junto com as mudanças na forma humana, ou seja, por análise da deformação da silhueta humana durante e após a queda, extraída a partir do método de subtração de fundo, como mostrado na Figura 5.



**Figura 5 – (a) Pessoa sendo monitorada; (b) Silhueta extraída do monitoramento a partir da subtração de fundo.**  
**Fonte: (WORRAKULPANIT *et al.*, 2014).**

A silhueta é então aproximada por uma elipse que informa sobre a forma e orientação da pessoa na imagem, como ilustrado na Figura 6.



**Figura 6 - Aproximação da silhueta humana por uma elipse.**  
**Fonte: adaptado de (WORRAKULPANIT *et al.*, 2014).**

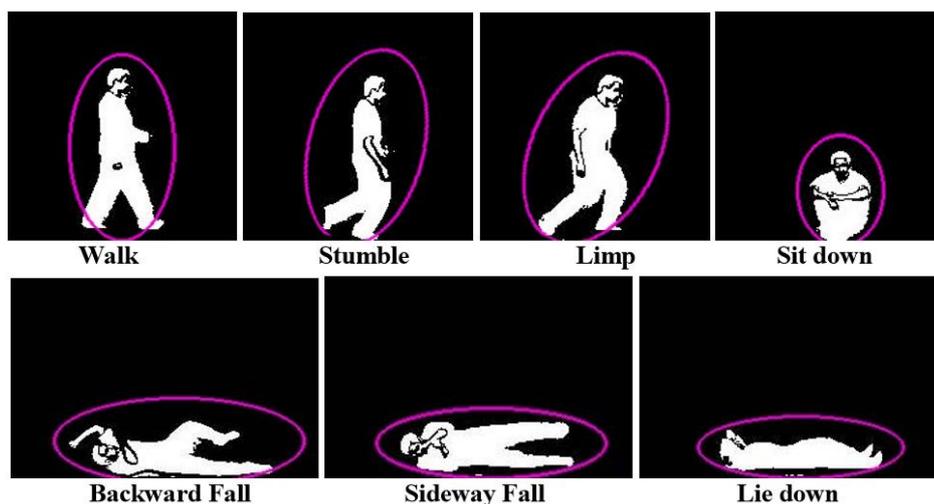
Para cada intervalo de tempo de 0,6 segundos é obtido o valor do coeficiente de quantidade de movimento  $C_{motion_i}$ , que a partir dos dados extraídos da elipse, mede entre 0% (nenhum movimento) e 100% (movimento total), depois é obtido o desvio padrão de  $C_{motion}$ , dado por:

$$\sigma_{C_{motion}} = \sqrt{\frac{1}{n} \sum_{i=1}^N (C_{motion_i} - \bar{c}_{motion})^2}. \quad (1)$$

Sendo  $N$  um número de *frames* especificados e  $\bar{c}_{motion}$  a média dos valores para os coeficientes  $C_{motion_i}$ , para o  $i$  variando de 1 até  $N$ . Se o valor do desvio padrão é elevado, então isso indica que pode existir queda, em seguida, é verificado se existe o movimento humano. Também é calculado o desvio padrão da orientação da elipse a partir do seu eixo e semi-eixo. Quando ocorre uma queda, há uma grande variação de movimento, isso acarreta uma mudança significativa no desvio padrão da orientação da elipse e conseqüentemente no valor de  $C_{motion}$ .

Foroughi *et al.* (2008) propuseram um sistema para detecção de queda humana em ambientes internos a partir de câmeras instaladas em posição de modo a obter um grande ângulo de visão. O método é baseado no melhor ajuste aproximado de uma elipse em torno do corpo humano, daí são extraídos vetores de característica para alimentar uma Rede Neural *MLP* de quatro camadas com esquema de aprendizagem *backpropagation* para classificação precisa dos movimentos e detecção

de uma possível queda (FOROUGHIL *et al.*, 2008). É usado o método de subtração de fundo para a segmentação do movimento e detectar a região que se desloca na imagem, subtraindo a imagem de fundo como referência da imagem atual, pixel a pixel. Os movimentos são analisados pela mudança na forma da elipse aproximada, mais precisamente na orientação e proporção, como ilustrado na Figura 7.



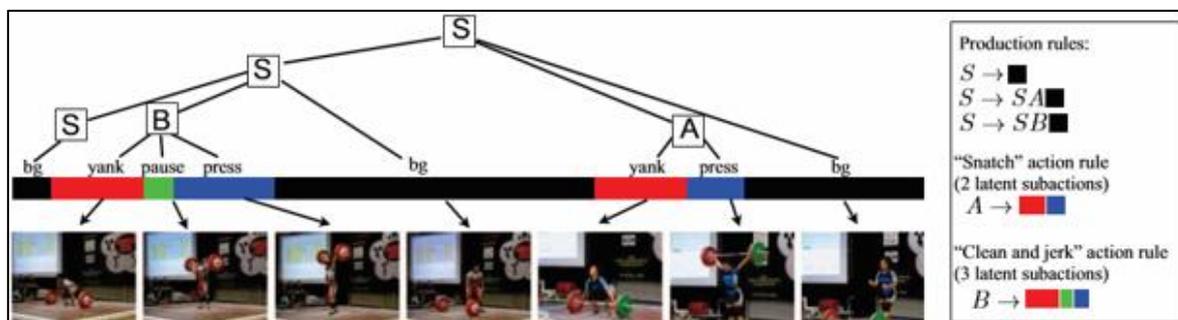
**Figura 7 - Elipse aproximada para diferentes posturas humana.**  
**Fonte: (FOROUGHIL *et al.*, 2008).**

Algumas metodologias para detecção de queda fazem uso de dispositivos automáticos, a exemplo dos acelerômetros, os quais detectam a magnitude e a direção da aceleração (TANG *et al.*, 2015; KANGAS *et al.*, 2008; KARANTONIS *et al.*, 2006). Outros baseiam-se em giroscópio (BOURKE *et al.*, 2008) que medem a orientação do corpo. Pisos detectores de vibração também são propostos para detectar a queda (ALWAN *et al.*, 2006).

Uma desvantagem dessas tecnologias é que estes sensores geralmente desgastam-se com facilidade e necessitam que as baterias sejam recarregadas regularmente para o funcionamento adequado, e para os pisos detectores de vibração a detecção da queda depende da dinâmica do chão, pois caso a queda seja de baixo impacto, não poderá ser detectada.

Não foi encontrado pesquisas utilizando gramáticas formais para detecção de queda, no entanto esta técnica é utilizada para reconhecer outros movimentos humanos, a exemplo da metodologia empregada por Pirsivash *et al.* (2014), que propuseram um algoritmo que analisa hierarquicamente fluxos de vídeos e faz uso de gramáticas para reconhecimento de ações relativas a

esportes olímpicos. Primeiramente foi realizada a segmentação de um vídeo de determinada ação. Em seguida são produzidas regras gramaticais que descrevem a segmentação realizada, e por fim a gramática é avaliada por uma máquina de estados finitos, em seguida emprega-se o *SVM* para classificar a ação que está sendo realizada. A metodologia proposta pelos autores divide uma determinada ação em várias sub-ações, por exemplo, a uma ação de “fazer chá” requer múltiplas sub-ações, tais como aquecimento da água, preparação da folha de chá, entre outras. A Figura 8 ilustra como foi realizada a segmentação e as regras de construção da gramática, para um determinado vídeo contendo a ação “levantamento de peso”. Esta ação pode ser composta de três sub-ações: arranque, pausa, sustentação. Ou de apenas duas sub-ações: arranque e sustentação. A construção das regras gramaticais para as duas sub-ações citadas são:  $S \rightarrow SB$  ou  $S \rightarrow SA$ , em que  $B$  indica três sub-ações e  $A$  indica duas sub-ações do vídeo segmentado,  $S$  é o antecedente da ação. Entretanto, a aplicação do trabalho destes autores restringe-se a alguns vídeos curtos de ações de atividades olímpicas. A abordagem proposta nesta pesquisa também realiza a segmentação e faz uso de gramática, no entanto, para reconhecer queda de idosos em ambientes internos.



**Figura 8 - Exemplo de segmentação e construção da gramática para um vídeo de levantamento de peso.**  
**Fonte: (PIRSIAVASH *et al.*, 2014).**

Os trabalhos citados que fazem uso do método da elipse aproximada para reconhecimento da queda não aplicam a gramática formal, e reconhecem a queda pela variação brusca do coeficiente da quantidade de movimento, que pode não ser tão confiável quando a queda não ocorre de modo brusco. Já a metodologia proposta neste trabalho além de também fazer uso da elipse aproximada, faz uso de gramática formal, sendo essa uma vantagem, pois reconhece a queda sendo ela brusca ou não.

## MONITORAMENTO SINTÁTICO DE MOVIMENTOS DE IDOSOS

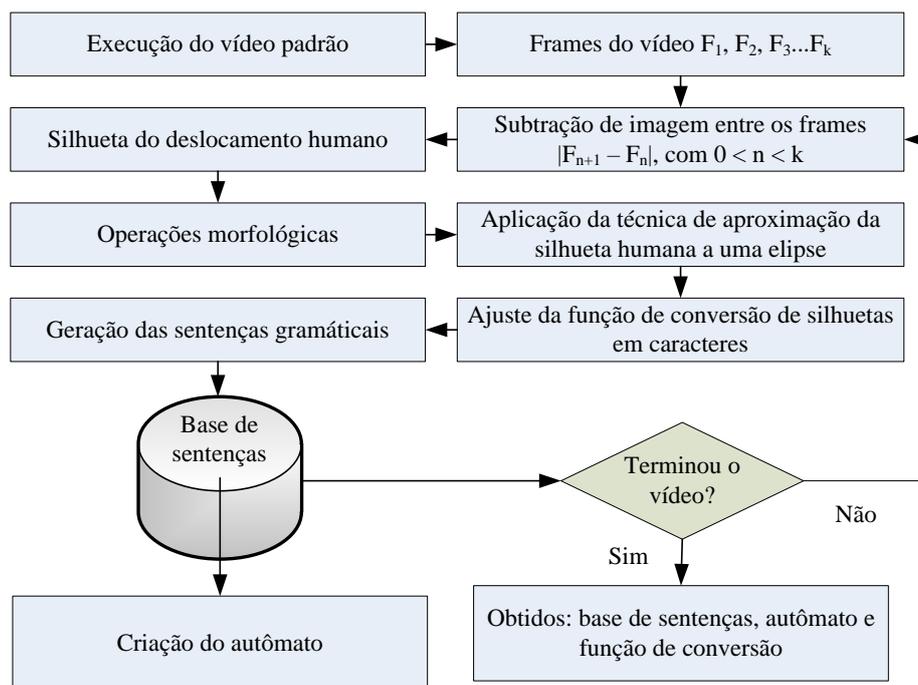
Este capítulo trata da abordagem proposta que viabiliza o monitoramento e detecção de queda de idosos em ambientes internos por meio de análise sintática de dados visuais. A abordagem é composta por dois processos principais:

1. Caracterização das gramáticas (treinamento);
2. Reconhecimento;

O primeiro processo corresponde ao treinamento mediante execução de um vídeo padrão da queda de um idoso, a partir do qual é caracterizada a gramática. O segundo processo realiza o reconhecimento propriamente dito.

### 3.1 FORMAÇÃO DAS SENTENÇAS GRAMATICAIS (TREINAMENTO)

A fase treinamento ocorre *off-line* do sistema e suas etapas estão ilustradas no fluxograma da Figura 9.



**Figura 9 - Fluxograma do processo de treinamento.**

**Fonte: Autor.**

De acordo com o Fluxograma da Figura 9, a partir da execução de um vídeo padrão, são gerados *frames* que são subtraídos entre si e geram silhuetas de deslocamentos do movimento humano. Nessas silhuetas são aplicadas *operações morfológicas* para aproximá-las a elipses. As inclinações das elipses são associadas a *zonas* de movimento, que por sua vez ajustam uma função de conversão que atribui um caractere para cada silhueta, deste modo ocorre caracterização das sentenças gramaticais.

Á partir da base de sentenças obtida é criado um Autômato Regular para reconhecer as palavras constituídas pelas cadeias de caracteres correspondentes aos movimentos e posições do idoso. Os estados do Autômato são referentes aos movimentos e posições e as regras de transições entre estes estados são as regras gramaticais estabelecidas.

As etapas da fase de treinamento serão posteriormente explicadas com detalhes nas seções seguintes.

### 3.2 RECONHECIMENTO

O segundo processo corresponde ao reconhecimento, que é baseado nas três etapas descritas a seguir e mostradas no fluxograma ilustrado na Figura 10:

- 1ª. Etapa: Obtenção dos dados visuais mediante monitoramento;
- 2ª. Etapa: Formação das palavras a partir das primitivas visuais extraídas;
- 3ª. Etapa: Reconhecimento da queda.

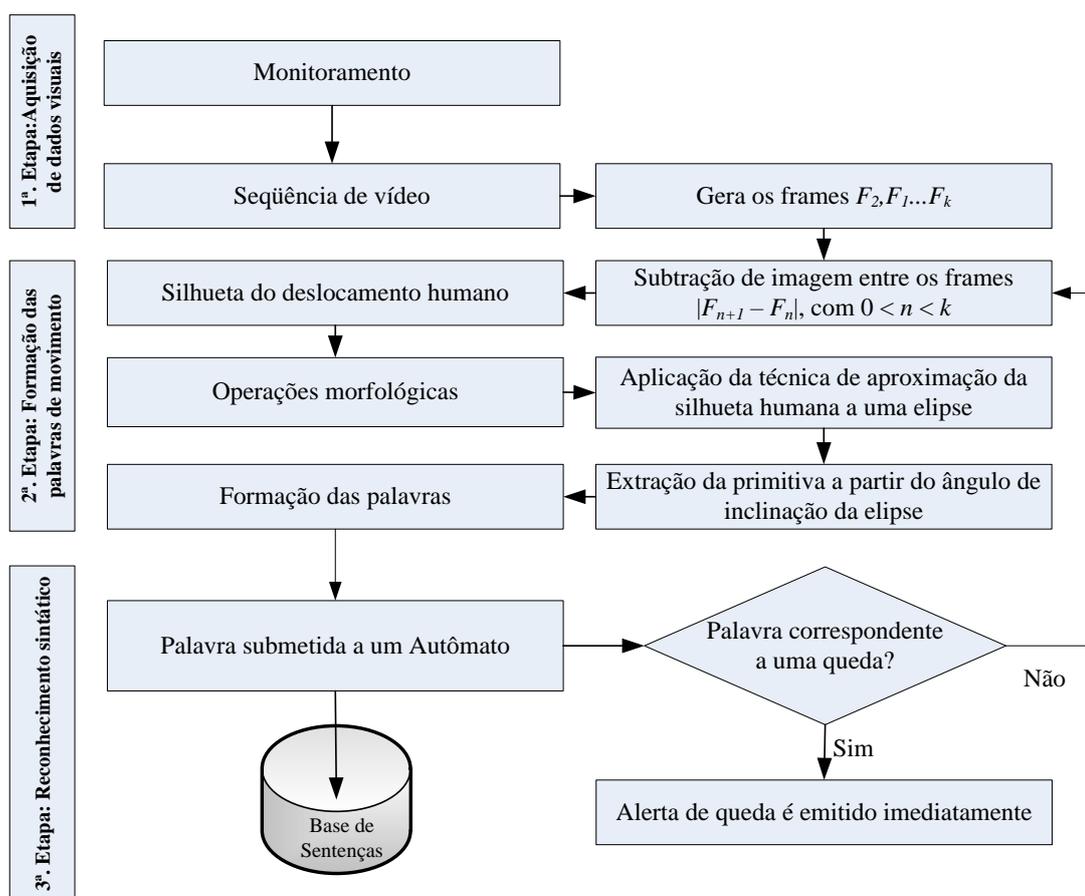


Figura 10 - Fluxograma do processo de reconhecimento.

Fonte: Autor.

De maneira semelhante ao treinamento, a partir de uma seqüência de vídeo são obtidos *frames* que são subtraídos gerando silhuetas do deslocamento humano. Nessas silhuetas são aplicadas *operações morfológicas* que irão aproximá-las a elipses. Nessas elipses são obtidas as primitivas a partir do ângulo de inclinação, os quais são convertidos em caracteres formando palavras de movimento. Um Autômato Regular, criado na fase de treinamento, realiza o reconhecimento a partir das regras gramaticais estabelecidas. Então caso a palavra constitua uma queda, o autômato

vai emitir um alerta. Caso contrário, o sistema continua a monitorar o movimento do idoso reconhecendo alguns movimentos ou possíveis desequilíbrios que possam ocasionar a queda.

As três etapas relacionadas ao processo de reconhecimento são explicadas detalhadamente a seguir:

### 3.2.1 1ª Etapa: Aquisição de dados visuais

Os dados de entrada são obtidos a partir de sequências de vídeo adquiridos por uma câmera convenientemente posicionada no ambiente interno, as imagens são obtidas continuamente até que a identificação do movimento ocorre por meio da subtração de imagem como descrita na seção 3.2.2. Se algum movimento é detectado a câmera continua capturando imagem e os *frames* são então processados convenientemente a partir das operações descritas na etapa a seguir.

### 3.2.2 2ª Etapa: Formação das palavras

A partir da sequência de imagens de entrada as palavras são processadas a partir das seguintes operações a seguir, para depois ser realizada a extração das primitivas e serem obtidos os caracteres para formação das palavras de movimento.

- **Subtração de imagem**

O objetivo da subtração de imagem é identificar a ocorrência de movimento e a obtenção de silhuetas do indivíduo. À medida que os *frames* são extraídos, cada frame seguinte é subtraído do anterior, resultando em uma silhueta de movimentação do corpo humano. A subtração de imagens consiste na detecção de mudança entre dois quadros de imagem  $f(x, y, t_i)$  e  $f(x, y, t_j)$ , mediante a comparação das duas imagens pixel a pixel, realizada pela equação:

$$d_{i,j}(x, y) = \begin{cases} 1 & \text{se } |f(x, y, t_i) - f(x, y, t_j)| > \gamma \\ 0 & \text{caso contrário} \end{cases} \quad (2)$$

Em que  $Y$  é um limiar predeterminado,  $x$  e  $y$  correspondem à posição cartesiana do pixel,  $t_i$  e  $t_j$  correspondem aos tempos em que as imagens foram obtidas (PICCARDI, 2004).

Como exemplo, temos que a Figura 11 (c) é o resultado da subtração entre a Figura 11 (a) e Figura 11 (b). Sendo a Figura 11 (a) e a Figura 11 (b) *frames* distintos e a Figura 11(c) a silhueta que representa o movimento entre eles. Em seguida são aplicadas as *operações morfológicas* na silhueta.



Figura 11 – (a) frame obtido em  $t_i$ ; (b) frame obtido em  $t_j$ ; (c) Resultado da subtração entre (a) e (b).  
Fonte: Autor.

- **Operações morfológicas aplicadas**

A função dos operadores morfológicos é tratar convenientemente a silhueta resultante da subtração de imagens  $d_{i,j}(x,y)$  de modo que seja possível aplicar o método da elipse aproximada.

A primeira operação morfológica a ser aplicada é a binarização, essa operação torna uma imagem em tons de cinza em uma imagem binária, ou seja, em preto e branco, para isto é considerado um parâmetro de limiar  $Y$  (GONZALEZ e WOODS, 2000). Neste caso, a imagem é mapeada pixel a pixel, comparando o valor de intensidade de nível de cinza com o valor de  $Y$ , este método é definido conforme a seguir:

$$b_{i,j}(x,y) = \begin{cases} 1 & \text{se } d_{i,j}(x,y) > Y \\ 0 & \text{caso contrário} \end{cases} \quad (3)$$

Sendo que, na equação (3),  $d_{i,j}(x,y)$  representa a função original resultante da diferença,  $b_{i,j}(x,y)$  a função binária e  $Y$  o valor do limiar. A Figura 12 corresponde a binarização da imagem ilustrada na Figura 11 (c).



**Figura 12 - Imagem binária da silhueta ilustrada na Figura 11 (c).**  
Fonte: Autor.

Após a binarização da imagem é necessário realizar a operação de fechamento para aproximá-la a uma elipse, por exemplo, o fechamento da imagem da Figura 12, resulta na imagem ilustrada na Figura 13. Soille (1999) define o fechamento  $\varphi$  de uma imagem  $f$  por um elemento estruturante  $B$  pela operação de dilatação de  $f$  seguido pela operação de erosão, com elemento estruturante transposto, determinado por:

$$\varphi(f) = \varepsilon \sim [\delta(f)]. \quad (4)$$

O elemento estruturante  $B$  pode ser uma cruz, quadrado, círculo, entre outros. A dilatação  $\delta$  de uma imagem por um elemento estruturante  $B$  é definida pela equação:

$$[\delta(f)](x) = \max f(x + B). \quad (5)$$

A erosão de  $f$  por um elemento estruturante  $B$  é definida pela equação:

$$[\varepsilon(f)](x) = \min\{f(x + B) - B\}. \quad (6)$$

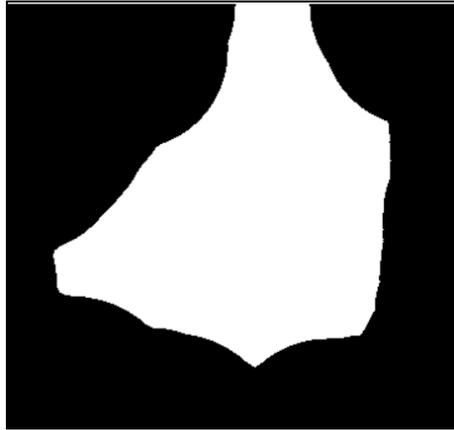


Figura 13 - Fechamento da imagem da silhueta da Figura 12.

Fonte: Autor.

- **Método da elipse aproximada**

O método da elipse aproximada é uma técnica que contorna a silhueta humana por uma elipse, e assim fornece informações sobre a orientação da posição de uma pessoa durante um determinado movimento. Essa técnica é possível fazendo o uso dos momentos espaciais e centrais da imagem da silhueta (ROUGIER *et al.*, 2007). Ou seja, a região de fechamento ilustrada na Figura 13, é representada por  $f(x,y)$  a partir da qual são obtidas as coordenadas centrais  $(\bar{x}, \bar{y})$  que definem o centro da elipse aproximada. A partir das coordenadas centrais são obtidos os momentos centrais de ordem zero, de primeira ordem e de segunda ordem, a partir dos quais é obtida a inclinação  $\theta$  da elipse aproximada. Logo as silhuetas de deslocamento geram regiões de fechamento que por sua vez geram elipses de contorno, como ilustrado na Figura 14.

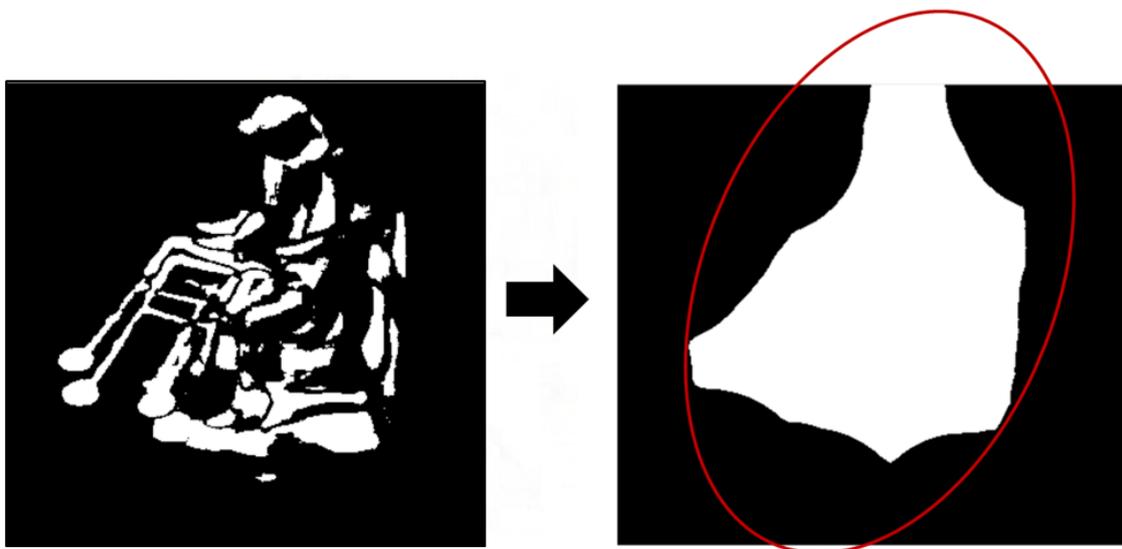


Figura 14 - Silhueta da forma humana gera uma região que é aproximada por uma elipse.  
Fonte: Autor.

A elipse aproximada é definida pelo seu centro  $(x, y)$ , ângulo teta ( $\theta$ ) que indica a orientação do seu eixo em relação ao eixo horizontal do sistema cartesiano, e pelos comprimentos  $a$  e  $b$  de seu eixo e semi-eixo respectivamente, como visto na Figura 15. A partir dela são extraídas as características da posição humana.

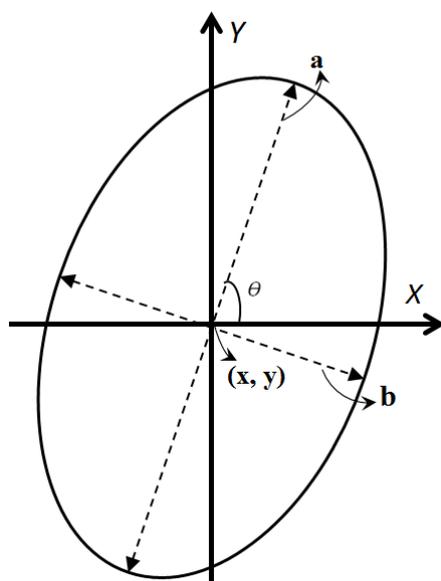


Figura 15 - Semi-eixo  $b$ , eixo maior  $a$ , centro  $(x, y)$  e inclinação  $\theta$  da elipse.  
Fonte: Autor.

Os momentos espaciais da região de fechamento  $f(x, y)$  são dados por:

$$m_{pq} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^p y^q f(x, y) dx dy; \quad \text{com: } p, q = 0, 1, 2, \dots \quad (7)$$

A partir dos momentos obtidos pela equação (7), calculam-se as coordenadas  $(\bar{x}, \bar{y})$  do centro de massa da imagem, pela relação entre os momentos espaciais de primeira ordem e o de ordem zero, dados por:

$$\bar{x} = m_{10}/m_{00}, \quad \bar{y} = m_{01}/m_{00}. \quad (8)$$

As coordenadas  $(\bar{x}, \bar{y})$  obtidas pela equação (8), definem então o centro da elipse aproximada, a partir do qual calculam-se os momentos centrais da seguinte maneira:

$$\mu_{pq} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - \bar{x})^p (y - \bar{y})^q f(x, y) dx dy. \quad (9)$$

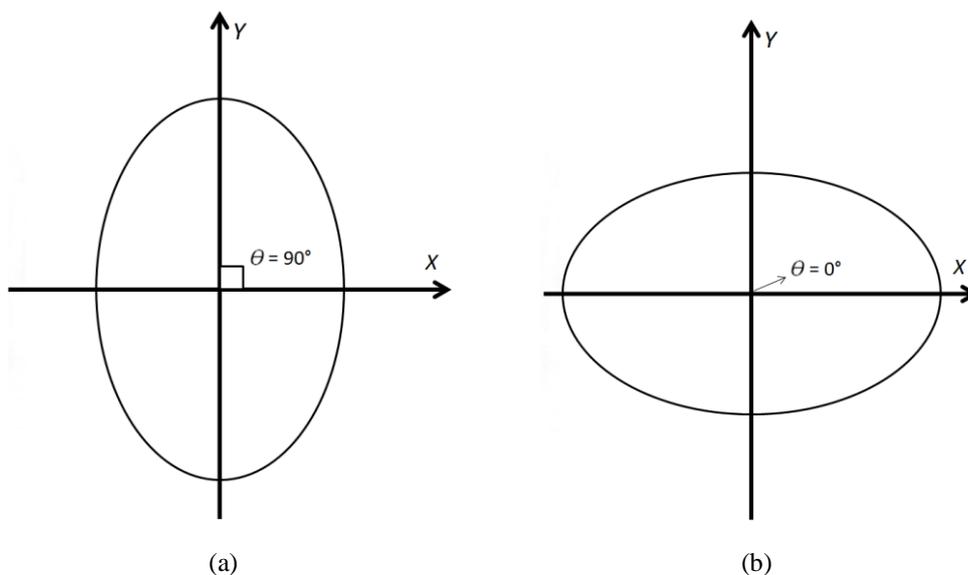
A partir dos momentos centrais de ordem zero, de primeira ordem e de segunda ordem obtidos pela equação (9) é calculado o ângulo de inclinação da elipse aproximada por:

$$\theta = \frac{1}{2} \arctan \left( \frac{2\mu_{11}}{\mu_{20} - \mu_{02}} \right). \quad (10)$$

Deste modo, após obter o centro  $(\bar{x}, \bar{y})$  e o ângulo de inclinação da elipse, a imagem de fechamento ilustrada na Figura 13, é então aproximada por uma elipse com inclinação  $\theta$ , semi-eixo  $b$  e eixo  $a$ .

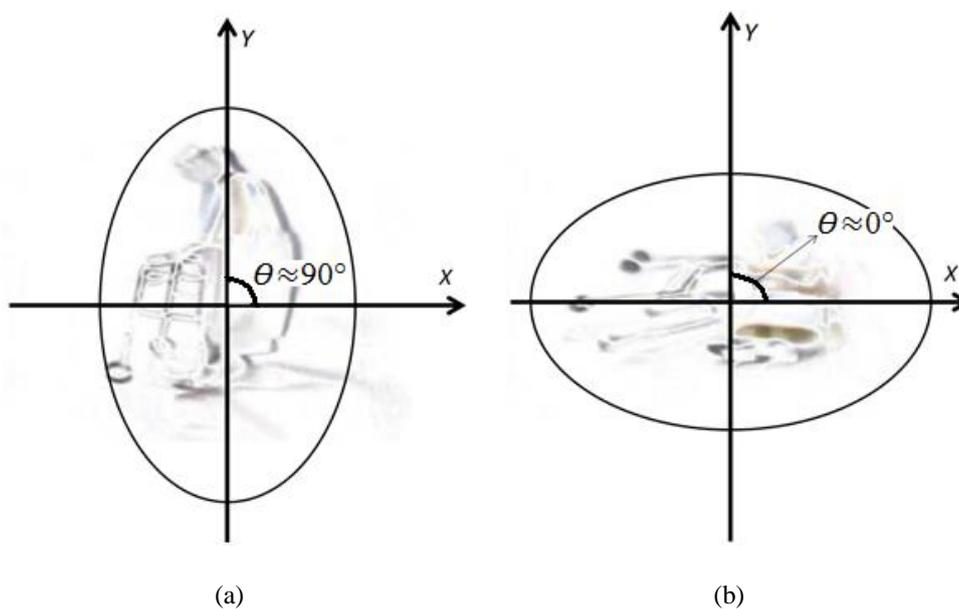
- **Zonas de posição**

Considerando a elipse aproximada obtida anteriormente, caso a inclinação  $\theta$  seja  $90^\circ$ , ela estará paralela ao eixo vertical  $Y$ , caso  $\theta$  seja  $0^\circ$ , a elipse estará paralela ao eixo horizontal  $X$ , do plano cartesiano  $X-Y$ , como mostrado na Figura 16.



**Figura 16 – (a) Elipse paralela ao eixo Y; (b) Elipse paralela ao eixo X.**  
**Fonte: Autor.**

Logo, temos como hipótese que a elipse que contorna a silhueta de uma pessoa andando possuirá inclinação próxima de  $90^\circ$ , e de uma pessoa caída próxima de  $0^\circ$ , como ilustra a Figura 17 (a) e (b) respectivamente.



**Figura 17 – Elipse aproximada da silhueta de uma pessoa andando (a) e caída (b).**  
**Fonte: Autor.**

Considerando essas hipóteses dadas anteriormente, o movimento humano é então parametrizado em três zonas de posição, *zona a* que corresponde a uma pessoa andando, *zona b* que

corresponde a uma *zona de transição* e *zona c* que corresponde a uma pessoa caída, como ilustra a Figura 18. As inclinações de cada zona são definidas durante o processo de treinamento.

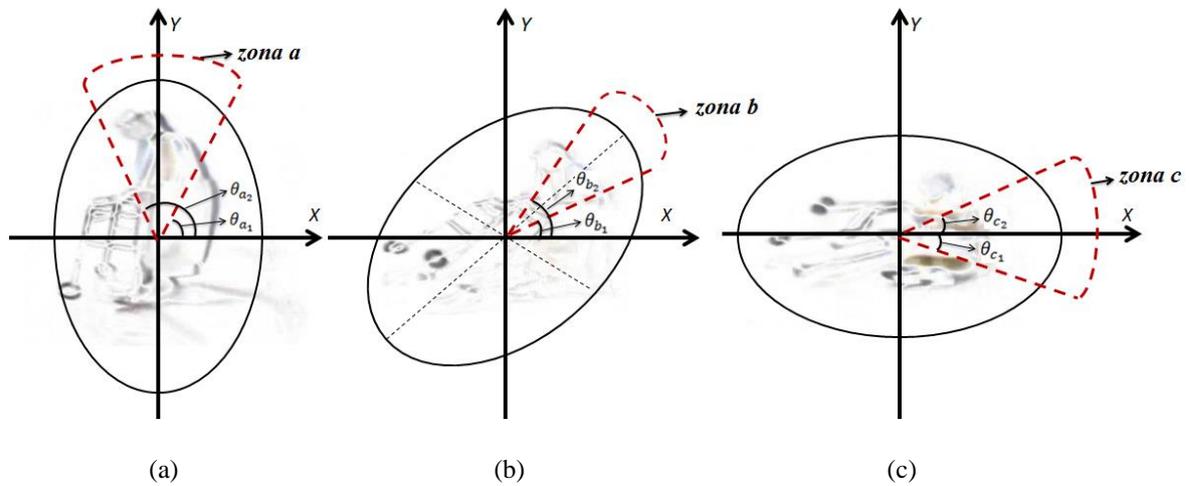


Figura 18 – Zonas de posição: (a) andando, (b) transição e (c) caído.

Fonte: Autor.

- **Extração de primitivas visuais**

Os módulos dos ângulos  $\theta_i$  das elipses de contorno podem assumir valores entre  $0^\circ$  e  $90^\circ$ , esse intervalo pode ser subdividido em regiões que representam o movimento da queda. Cada ângulo tomado dentro de uma região específica constitui a inclinação da silhueta do movimento humano que representa uma primitiva das quais são obtidos os caracteres como detalhado a seguir.

- **Obtenção dos caracteres a partir das primitivas e função de conversão**

A função de conversão  $f(\theta)$  para obter os caracteres correspondentes às zonas de posição da Figura 24 é definida por:

$$f(\theta_i) = \begin{cases} a, & \text{se } \theta_{a_1} < \theta_i \leq \theta_{a_2} \\ b, & \text{se } \theta_{b_1} < \theta_i \leq \theta_{b_2} \\ c, & \text{se } \theta_{c_1} < \theta_i \leq \theta_{c_2} \\ f(\theta_{i-1}), & \text{se } \theta_i = \text{indeterminado} \end{cases} \quad (11)$$

Os ângulos  $(\theta_{a_1}, \theta_{a_2}, \theta_{b_1}, \theta_{b_2}, \theta_{c_1}, \theta_{c_2})$  são definidos durante a fase de treinamento e determinam as zonas posições. Ou seja, quando a silhueta possui o ângulo de inclinação  $\theta_i$  entre

$\theta_{a_1}$  e  $\theta_{a_2}$  a primitiva visual recebe o caractere  $a$ . Quando a silhueta possui o ângulo de inclinação  $\theta_i$  entre  $\theta_{b_1}$  e  $\theta_{b_2}$  a primitiva visual recebe o caractere  $b$ , e quando a silhueta possui o ângulo de inclinação  $\theta_i$  entre  $\theta_{c_1}$  e  $\theta_{c_2}$  a primitiva visual recebe o caractere  $c$ . Deste modo, o movimento humano é descrito por cadeias de caracteres constituindo palavras de movimento.

Quando  $\theta_i$  é indeterminado, então a primitiva visual recebe o mesmo caractere obtido anteriormente, ou seja,  $f(\theta_i) = f(\theta_{i-1})$  para considerar a posição atual igual à posição anterior, pois em alguns casos a oclusão pode gerar silhuetas com mais de uma região de fechamento, deste modo o ângulo de inclinação da silhueta é indeterminado, pois, o método da elipse aproximada aplica-se em silhuetas com apenas uma região de fechamento.

- **Formação das regras Gramaticais**

As cadeias de caracteres constituem milhares de palavras. Dentre elas algumas são de interesse, ou seja, as que representam posição ou movimento do idoso.

São necessárias aplicações de regras gramaticais tanto para formação das palavras quanto para reconhecê-las sintaticamente.

Estas regras são dadas a seguir:

- Um caractere obtido no tempo  $t$  representa uma primitiva visual também obtida no tempo  $t$ ;
- As palavras devem ser constituídas por caracteres obtidos hierarquicamente no tempo;
- As palavras de movimento/posição do idoso são definidas durante a fase de treinamento;
- A gramática é definida pela quádrupla

$$G = \langle V, \Sigma, R, S \rangle.$$

Em que:

- $\Sigma = \{a, b, c\}$ , constitui o alfabeto;
- $V = (S, \Sigma, T)$ , sendo  $S$  e  $T$  as variáveis de partida e terminal, respectivamente;
- $R$  constitui as regras que são dadas por:
  - $S \rightarrow aS$

- ii.  $S \rightarrow bS$
- iii.  $S \rightarrow cS$
- iv.  $S \rightarrow aT$
- v.  $S \rightarrow bT$
- vi.  $S \rightarrow cT$

### 3.2.3 Reconhecimento sintático

As regras gramaticais constituem então um Autômato Finito Determinístico–*AFD* que realiza o reconhecimento das palavras. O Autômato é definido por:  $A = \langle Q, \Sigma, \delta, S, T \rangle$ , sendo:

- a)  $Q = \{q_a, q_b, q_c\}$ , os estados que podem ser alcançados, em que:
  - a.  $q_a$  é o estado associado ao movimento andando;
  - b.  $q_b$  é o estado associado ao movimento de transições entre os eventos (desequilíbrio);
  - c.  $q_c$  é o estado associado a posição caído;
- b)  $\Sigma = \{a, b, c\}$ , o alfabeto;
- c)  $\delta$ , função de transição de estado de acordo com as regras gramaticais que foram definidas;
- d)  $S, T$ , estados de partida e terminal, respectivamente.

Considerando essas regras e definições é possível a criar um *AFD* a partir da execução de um vídeo padrão da queda de um idoso, como é mostrado no capítulo seguinte.

---

## APLICAÇÕES EXPERIMENTAIS

As aplicações experimentais e seus respectivos resultados, obtidos a partir da metodologia proposta serão apresentados neste capítulo.

### 4.1 PROCEDIMENTOS E MÉTODOS

As aplicações foram realizadas para cenas reais de monitoramento de idosos que sofreram queda, obtidas em (ROBINOVITCH *et al.*, 2012), e também para simulações de queda durante monitoramentos realizados no ICOMP da Universidade Federal do Amazonas – UFAM. Também foram realizadas aplicações para comparar a metodologia proposta com outros métodos para reconhecimento de queda.

O método proposto foi implementado no *Matlab*® (*The MathWork, Inc. Software*) versão 2010, utilizando um computador de 64 *bits* com processador Intel Core i5 de 2,20 Ghz e 8 GBytes de memória RAM em ambiente Windows 10. Foi utilizado o software *KMPlayer* para gerar os *frames*.

#### 4.1.1 Métricas

O modelo é avaliado em relação ao processo de classificação por meio das seguintes métricas:

- **Acurácia**

A *acurácia* mede a proporção de predições corretas, sem levar em consideração o que é positivo e o que é negativo. Esta medida é altamente suscetível a desbalanceamentos do conjunto de dados e pode facilmente induzir a uma conclusão errada sobre o desempenho do sistema, é medida pelo total de acertos dividido pelo total de dados no conjunto, expressa por:

$$acurácia = \frac{(VP + VN)}{(VP + VN + FP + FN)}. \quad (12)$$

- **Sensibilidade**

A *sensibilidade* mede a proporção de verdadeiros positivos, ou seja, a capacidade do sistema em prever corretamente a condição para casos que realmente a têm, é medida pelos acertos positivos dividido pelo total de positivos, expressa por:

$$sensibilidade = \frac{VP}{(VP + FN)}. \quad (13)$$

- **Especificidade**

A *especificidade* mede a proporção de verdadeiros negativos, ou seja, a capacidade do sistema em prever corretamente a ausência da condição para casos que realmente não a têm, é calculada pelos acertos negativos dividido pelo total negativo, expressa por:

$$especificidade = \frac{VN}{(VN + FP)}. \quad (14)$$

- **Eficiência**

A *eficiência* é a média aritmética da *Sensibilidade* e *Especificidade*. Na prática, a *sensibilidade* e a *especificidade* variam em direções opostas. Isto é, geralmente, quando um método é muito sensível a positivos, tende a gerar muitos falso-positivos, e vice-versa. Assim, um método de decisão perfeito (100% de *sensibilidade* e 100% *especificidade*) raramente é alcançado, e um balanço entre ambos deve ser atingido, é expressa por:

$$eficiência = \frac{(sensibilidade + especificidade)}{2}. \quad (15)$$

- **Valor Preditivo Positivo (VPP)**

O *VPP* mede a proporção de verdadeiros positivos em relação a todas as predições positivas. Esta medida é altamente suscetível a desbalanceamentos do conjunto de dados e pode facilmente induzir a uma conclusão errada sobre o desempenho do sistema. É calculado pelo total de acertos positivos dividido pelo total de predições positivas, é obtido por:

$$VPP = \frac{VP}{(VP + FP)}. \quad (16)$$

- **Valor Preditivo Negativo (VPN)**

A *VPN* mede a proporção de verdadeiros negativos em relação a todas as predições negativas. Esta medida é altamente suscetível a desbalanceamentos do conjunto de dados e pode facilmente induzir a uma conclusão errada sobre o desempenho do sistema. É calculado pelos acertos negativos divididos pelo total de predições negativas, como mostrado a seguir:

$$VPN = \frac{VN}{(VN + FN)}. \quad (17)$$

- **Coefficiente de Correlação de Matthews – Coeficiente (*phi*)**

O coeficiente de correlação de Matthews é uma medida de qualidade de duas classificações binárias que pode ser usada mesmo se as classes possuem tamanhos bastante diferentes. Retorna um valor entre (-1) e (+1), em que um coeficiente de (+1) representa uma predição perfeita, (0) representa uma predição aleatória média, e (-1) uma predição inversa. Esta estatística é equivalente ao coeficiente *phi*, e tenta, assim como a *eficiência*, resumir a qualidade da tabela de contingência em um único valor numérico passível de ser comparado. Nota-se que, se qualquer uma das somas no denominador for igual a (0), o denominador pode ser considerado (1), resultando em um *phi* igual a (0) que seria o limite correto para esta situação, é expresso por:

$$phi = \frac{(VP * VN - FP * FN)}{\sqrt{(VP + FP) * (VP + FN) * (VN + FP) * (VN + FN)}}. \quad (18)$$

## 4.2 PLANEJAMENTO EXPERIMENTAL

As aplicações foram divididas em três grupos de experimentos: treinamento, aplicações e comparação com outros métodos. Estes grupos estão relacionadas com as atividades de risco em que o idoso pode se deparar, como mostradas na Tabela 1.

Tabela 1 – Procedimento Experimental.

ETAPAS DO EXPERIMENTO		
GRUPOS	MONITORAMENTO	DETALHE
TREINAMENTO	Idoso cai ao deambular	Vídeo padrão em que um idoso está deambulando em sua residência e sofre queda.
APLICAÇÕES	1: Idoso derrubado pelo seu cão	Idoso está caminhando com o seu cão, sendo derrubado pelo mesmo.
	2: Idoso cai ao realizar exercício	Idoso está realizando exercício em sua casa e acaba tropeçando e sofrendo queda.
	3: Idoso cai ao tropeçar em objeto	Idoso está caminhando no corredor de um prédio, tropeça em um carrinho de limpeza e cai.
	4: Idoso cai ao levantar da cadeira	Idoso está sentado à mesa e ao sair da cadeira acaba sofrendo uma queda.
	5: Idoso cai ao descer escadaria	Idoso descer a escadaria de um prédio durante a noite tropeça e acaba sofrendo queda.
COMPARAÇÃO COM OUTROS MÉTODOS	Pessoa está caminhando e sofre queda	Dados sintéticos usados na aplicação de Foroughi <i>et al.</i> (2008) com Redes Neural MLP.

### 4.2.1 Primeira etapa: realização do treinamento e definição da base de regras:

O treinamento foi realizado com um vídeo padrão do monitoramento de um idoso em sua residência. O vídeo do monitoramento mostra o idoso realizando uma atividade de risco que é deambular sozinho na residência, de repente sofre uma queda. O vídeo do monitoramento está disponível *on-line*<sup>1</sup>, cuja sequência de *frames* está ilustrada na Figura 19, enumerados na sequência de F1 a F50.

<sup>1</sup> <https://www.youtube.com/watch?v=p5i4z3sNaKM>



**Figura 19 - Frames da queda de um idoso que estava deambulando.**  
Fonte: Autor.

Cada *frame* da Figura 19 é subtraído do frame anterior, ou seja,  $|F_n - F_{n-1}|$ , gerando uma silhueta da pose do corpo humano que após aplicação das *operações morfológicas* e do método da elipse aproximada, são obtidos seus respectivos ângulos de inclinação  $\theta_i$ , como ilustra a Figura 20.

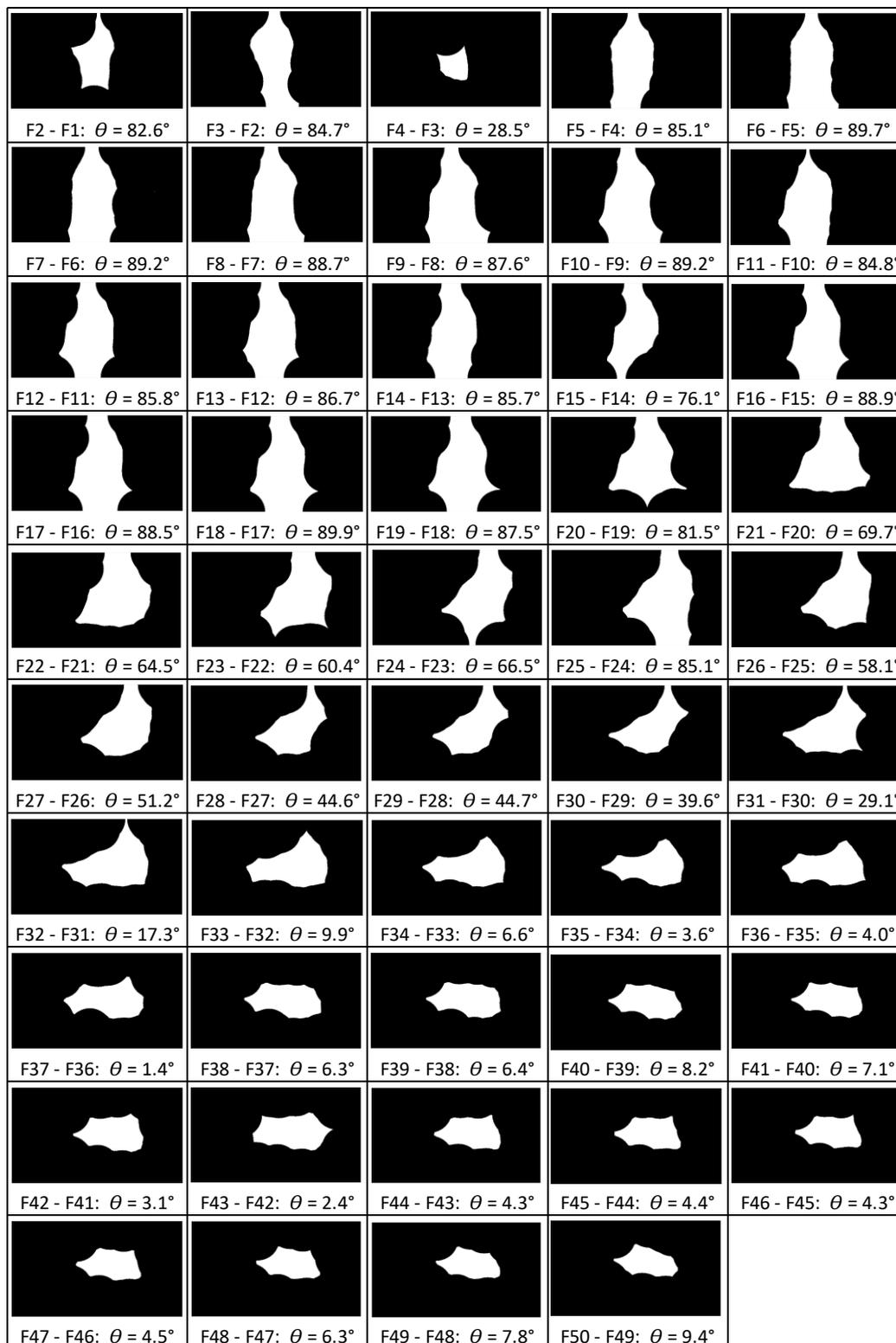


Figura 20 – Silhuetas das subtrações dos *frames* da Figura 19 e suas respectivas inclinações.

Fonte: Autor.

Observa-se na Figura 20 que as primeiras silhuetas ilustram o contorno da forma humana na vertical, condizendo com o deslocamento do movimento do idoso andando que também é na vertical, conseqüentemente a elipse que contorna a silhueta tem seu eixo próximo do eixo vertical com valor da inclinação  $\theta$  próximo de  $90^\circ$ . Progressivamente as silhuetas vão aproximando-se da horizontal condizendo com o movimento da queda e conseqüentemente  $\theta$  vai aproximando-se de  $0^\circ$ , até que as silhuetas permanecem quase que invariantes na horizontal com  $\theta$  muito aproximado a  $0^\circ$  caracterizando a fase pós-queda. Observa-se que a subtração dos *frames* ( $F_{51} - F_{50}$ ) não resulta em nenhuma silhueta, isto porque a partir deste instante o idoso permaneceu imóvel e não ocorreu movimento entre estes dois *frames*, neste caso o ângulo  $\theta$  é considerado o mesmo da silhueta gerada pela subtração anterior, ou seja  $\theta_{(F_{51}-F_{50})} = \theta_{(F_{50}-F_{49})}$  de acordo com a equação (13).

Foram obtidos os valores das inclinações ( $\theta_{a_1}$ ,  $\theta_{a_2}$ ,  $\theta_{b_1}$ ,  $\theta_{b_2}$ ,  $\theta_{c_1}$ ,  $\theta_{c_2}$ ) que definem as zonas de posição. Esses resultados estão ilustrados na Tabela 2.

**Tabela 2 - Parâmetros obtidos da fase de treinamento**

ZONAS DE POSIÇÃO	$\theta_1 < \theta_i \leq \theta_2$	CARACTERES ASSOCIADOS
<i>zona a</i>	$55^\circ < \theta_i \leq 125^\circ$	<i>a</i>
<i>zona b</i>	$35^\circ < \theta_i \leq 55^\circ$	<i>b</i>
<i>zona c</i>	$-35^\circ < \theta_i \leq 35^\circ$	<i>c</i>

A partir da Tabela 2 observa-se que as *zonas a* e *c* possuem maiores ângulos de abrangência em relação à *zona b*, devido o fato dos movimentos *andando* e *caído* serem geralmente mais duradouros que a *transição*, como pode ser constatado na Figura 19.

A Tabela 3 exhibe a seqüência de caracteres obtidos para cada silhueta da Figura 20. Cada um deles resulta da aplicação da função de conversão expressa em (11).



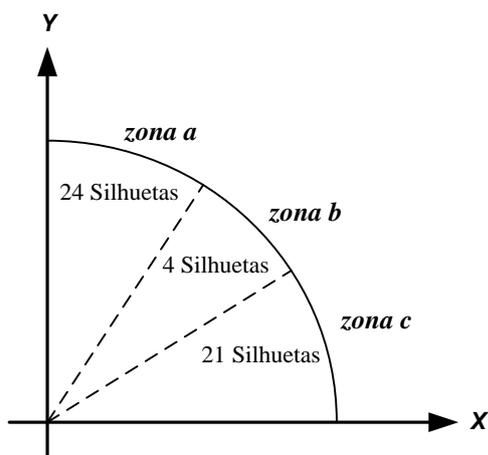


Figura 22 – Distribuição das silhuetas por zonas de posição durante o treinamento.

Fonte: Autor.

A Tabela 4 mostra as palavras de posição obtidas durante o treinamento, descritas por meio de expressões regulares.

Tabela 4 - Palavras que representam as Posições e Movimento.

CARACTERES	PALAVRA	POSIÇÃO/MOVIMENTO
$a^*$	$aaa...a$	Andando/deambulando
$b^*$	$bbb...b$	Desequilíbrio ou transição
$c^*$	$ccc...c$	Caído
$a^*b^*$	$(aaa...a)(bbb...b)$	Andando → desequilibra
$a^*c^*$	$(aaa...a)(ccc...c)$	Andando → cai
$b^*a^*$	$(bbb...b)(aaa...a)$	Desequilibra → andando
$b^*c^*$	$(bbb...b)(ccc...c)$	Desequilibra → cai
$c^*b^*$	$(ccc...c)(bbb...b)$	Caído → desequilibra
$c^*a^*$	$(ccc...c)(aaa...a)$	Caído → andando

A partir das expressões regulares da Tabela 4 é criado um Autômato Regular  $A = (Q, \Sigma, \delta, S, T)$ , para realizar o reconhecimento das palavras que representam as posições e movimentos do idoso, como é ilustrado na Figura 23.

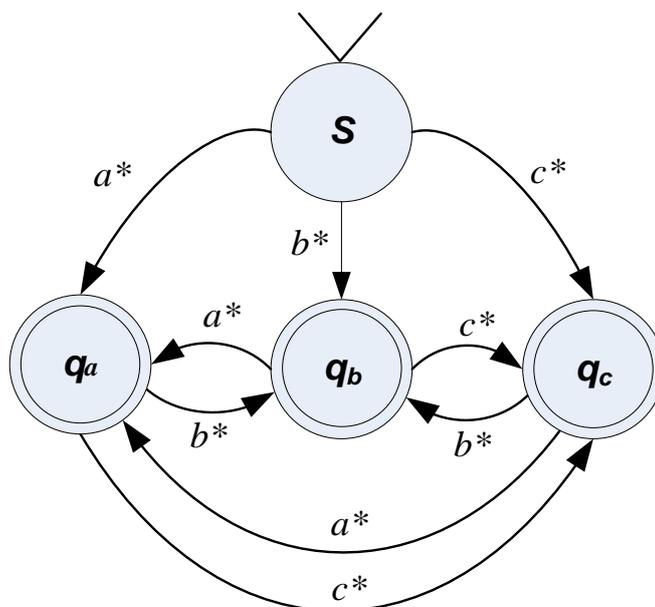


Figura 23 - Autômato  $A = (Q, \Sigma, \delta, S, T)$ .

Fonte: Autor.

De acordo com o Autômato da Figura 23, uma sequência de caracteres  $a$  leva ao estado  $q_a$  correspondente a posição *andando*, uma sequência de caracteres  $b$  resulta no estado  $q_b$  que corresponde à posição de *transição*, já uma sequência de caracteres  $c$  leva ao estado  $q_c$  que corresponde ao estado *caído*. O Autômato pode migrar para quaisquer dos estados dependendo da cadeia de caracteres obtida.

#### 4.2.2 Aplicação 01: Idoso é derrubado pelo seu cão

Esta aplicação foi realizada durante o monitoramento do corredor de um prédio e flagra o momento em que um casal de idosos cruza o corredor com seu cachorro que derruba um deles. O vídeo está disponível *on-line*<sup>2</sup>.

A sequência de *frames* que representam o monitoramento durante a queda está ilustrada na Figura 24, enumerados na sequência de  $F_1$  a  $F_{50}$ .

<sup>2</sup> [https://www.youtube.com/watch?v=1IsM08Sh\\_wg](https://www.youtube.com/watch?v=1IsM08Sh_wg)

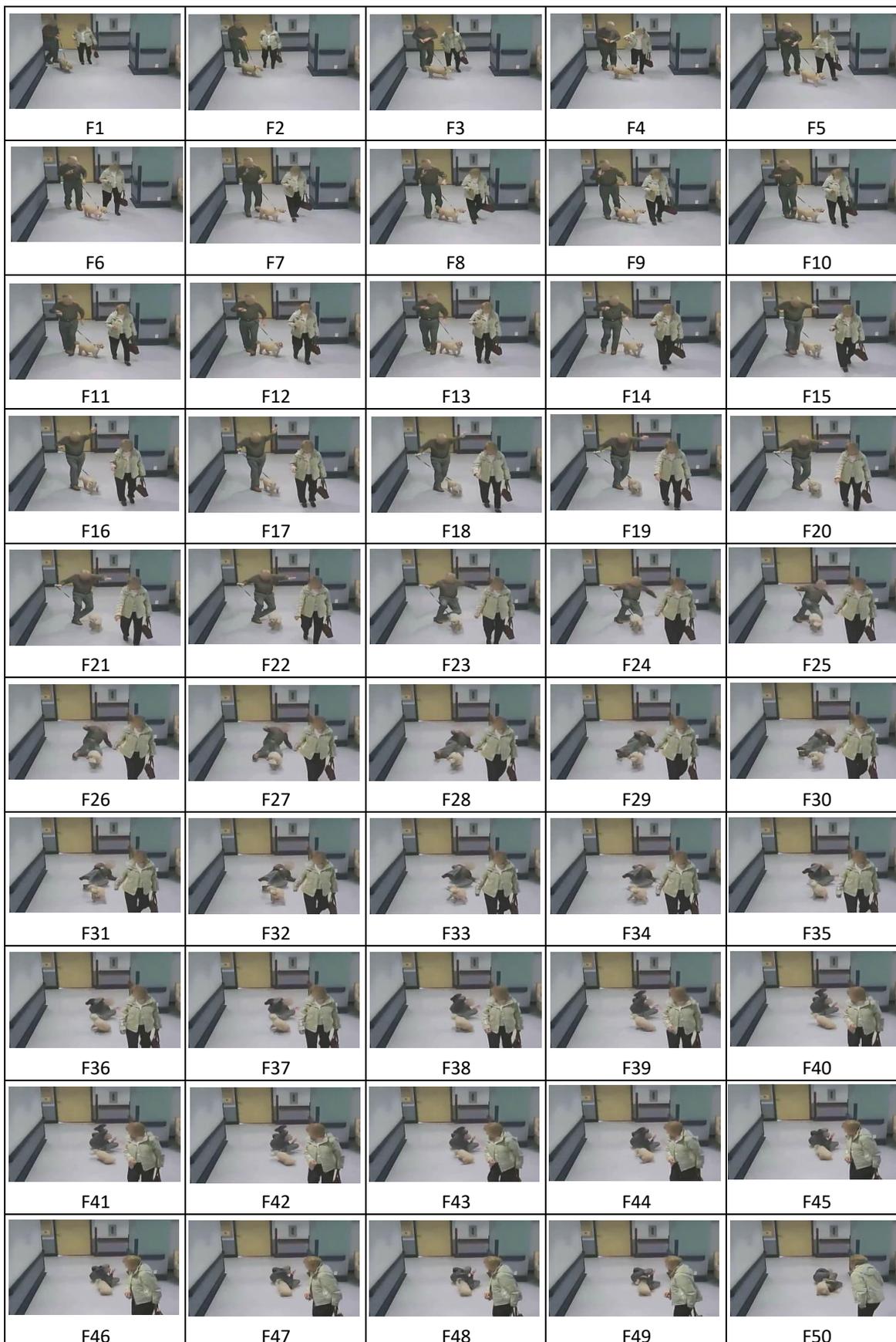


Figura 24 - Frames da queda de um idoso derrubado pelo seu cão.

Fonte: Autor.

As silhuetas de deslocamento do corpo humano resultante de cada subtração *frame a frame*, e os respectivos ângulos de inclinação  $\theta$  das elipses aproximadas, estão ilustrados na Figura 25.

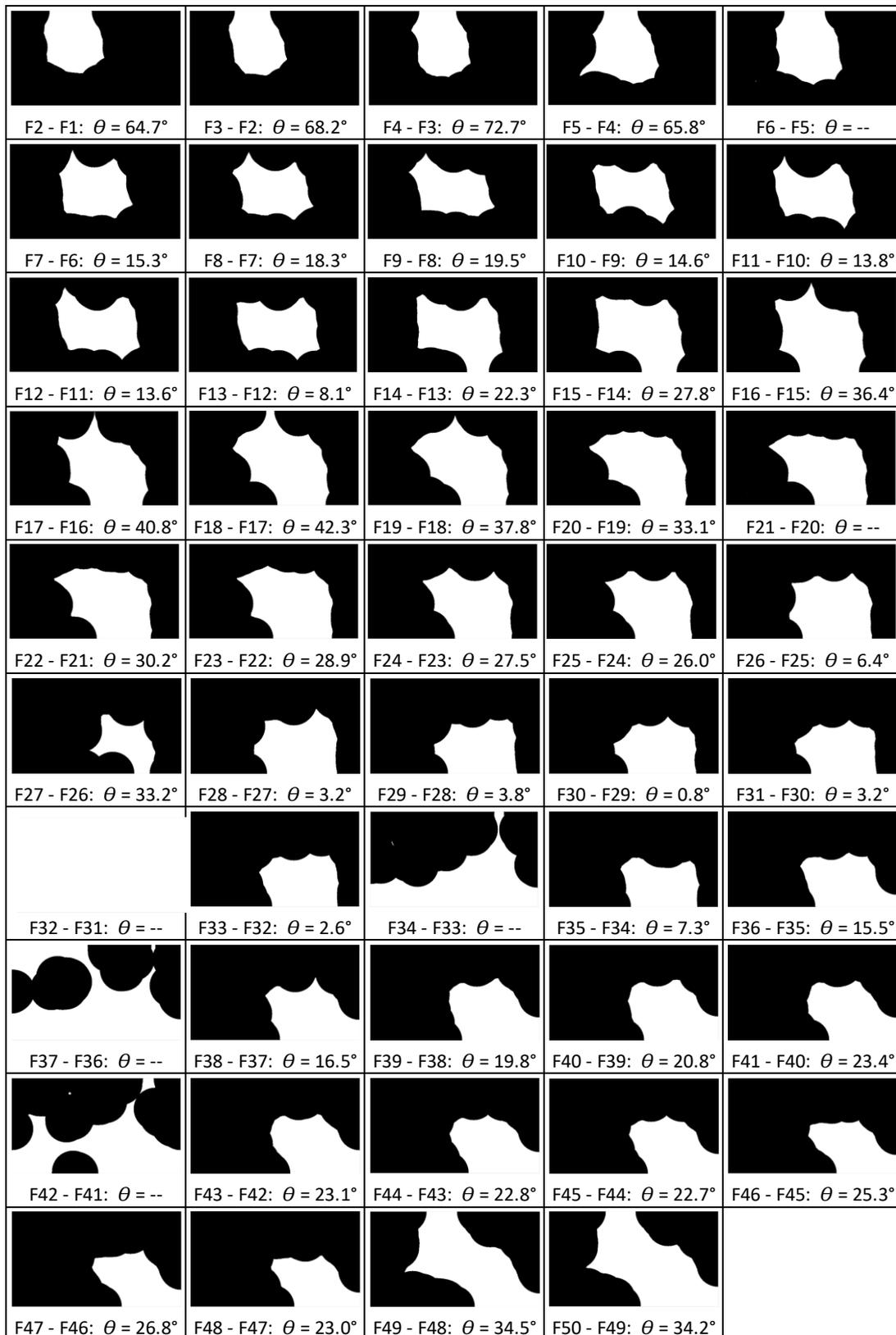


Figura 25 - Silhuetas das subtrações dos *frames* da Figura 24 e suas respectivas inclinações.

Fonte: Autor.

A Tabela 5 exibe os caracteres obtidos para cada silhueta da Figura 25, de acordo com a função de conversão expressa na equação (11).

Tabela 5 – Caracteres obtidos da aplicação 01.

SEQUÊNCIA DE CARACTERES OBTIDOS				
1: $f(64,7^\circ) = a$	2: $f(68,2^\circ) = a$	3: $f(72,7^\circ) = a$	4: $f(65,8^\circ) = a$	5: -- = $a$
6: $f(15,3^\circ) = c$	7: $f(18,3^\circ) = c$	8: $f(19,5^\circ) = c$	9: $f(14,6^\circ) = c$	10: $f(13,8^\circ) = c$
11: $f(13,6^\circ) = c$	12: $f(8,1^\circ) = c$	13: $f(22,3^\circ) = c$	14: $f(27,8^\circ) = c$	15: $f(36,4^\circ) = b$
16: $f(40,8^\circ) = b$	17: $f(42,3^\circ) = b$	18: $f(37,8^\circ) = b$	19: $f(33,1^\circ) = b$	20: -- = $b$
21: $f(30,2^\circ) = c$	22: $f(28,9^\circ) = c$	23: $f(27,5^\circ) = c$	24: $f(26,0^\circ) = c$	25: $f(6,4^\circ) = c$
26: $f(33,2^\circ) = c$	27: $f(3,2^\circ) = c$	28: $f(3,8^\circ) = c$	29: $f(0,8^\circ) = c$	30: $f(3,2^\circ) = c$
31: -- = $c$	32: $f(2,6^\circ) = c$	33: -- = $c$	34: $f(7,3^\circ) = c$	35: $f(15,5^\circ) = c$
36: -- = $c$	37: $f(16,5^\circ) = c$	38: $f(19,8^\circ) = c$	39: $f(20,8^\circ) = c$	40: $f(23,4^\circ) = c$
41: -- = $c$	42: $f(23,1^\circ) = c$	43: $f(22,8^\circ) = c$	44: $f(22,7^\circ) = c$	45: $f(25,3^\circ) = c$
46: $f(26,8^\circ) = c$	47: $f(23,0^\circ) = c$	48: $f(34,5^\circ) = c$	49: $f(34,2^\circ) = c$	

Para algumas silhuetas o ângulo de inclinação não pode ser determinado, devido o fato destas silhuetas serem compostas por mais de uma região de contorno, ocasionadas pelo movimento conjunto da outra pessoa e do cachorro durante a queda do idoso ou pela subtração não resultar em nenhuma silhueta, logo, atribui-se aos seus caracteres correspondentes o mesmo caractere anterior, como mostrado nas posições 5, 20, 31, 33, 36 e 41 da Tabela 5.

O número de *frames* obtidos em cada zona de posição durante a aplicação 01 está ilustrado na Figura 26. A *zonas a* e *zona b* são compostas por 5 e 6 *frames* respectivamente, já a *zona c* é composta por 38 *frames* por corresponder a maior duração do vídeo, sendo a queda do idoso.

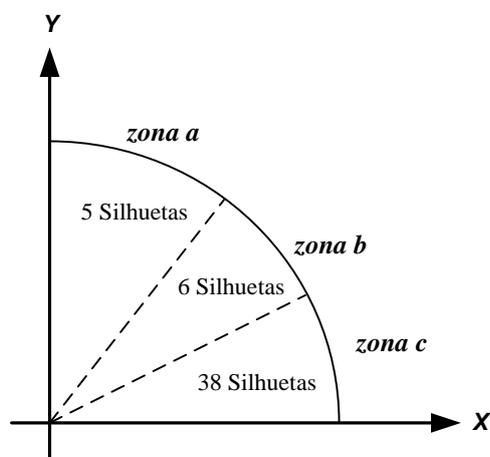


Figura 26 – Distribuição das silhuetas por zonas durante a aplicação 01.

Fonte: Autor.



tropeçar em seu calcanhar, resultando em uma queda para trás. Durante a queda ele colide parcialmente com uma árvore de natal. O vídeo do monitoramento está disponível *on-line*<sup>3</sup>.

Os *frames* da queda do idoso estão ilustrados na Figura 28, enumerados na sequência de  $F_1$  a  $F_{50}$ .

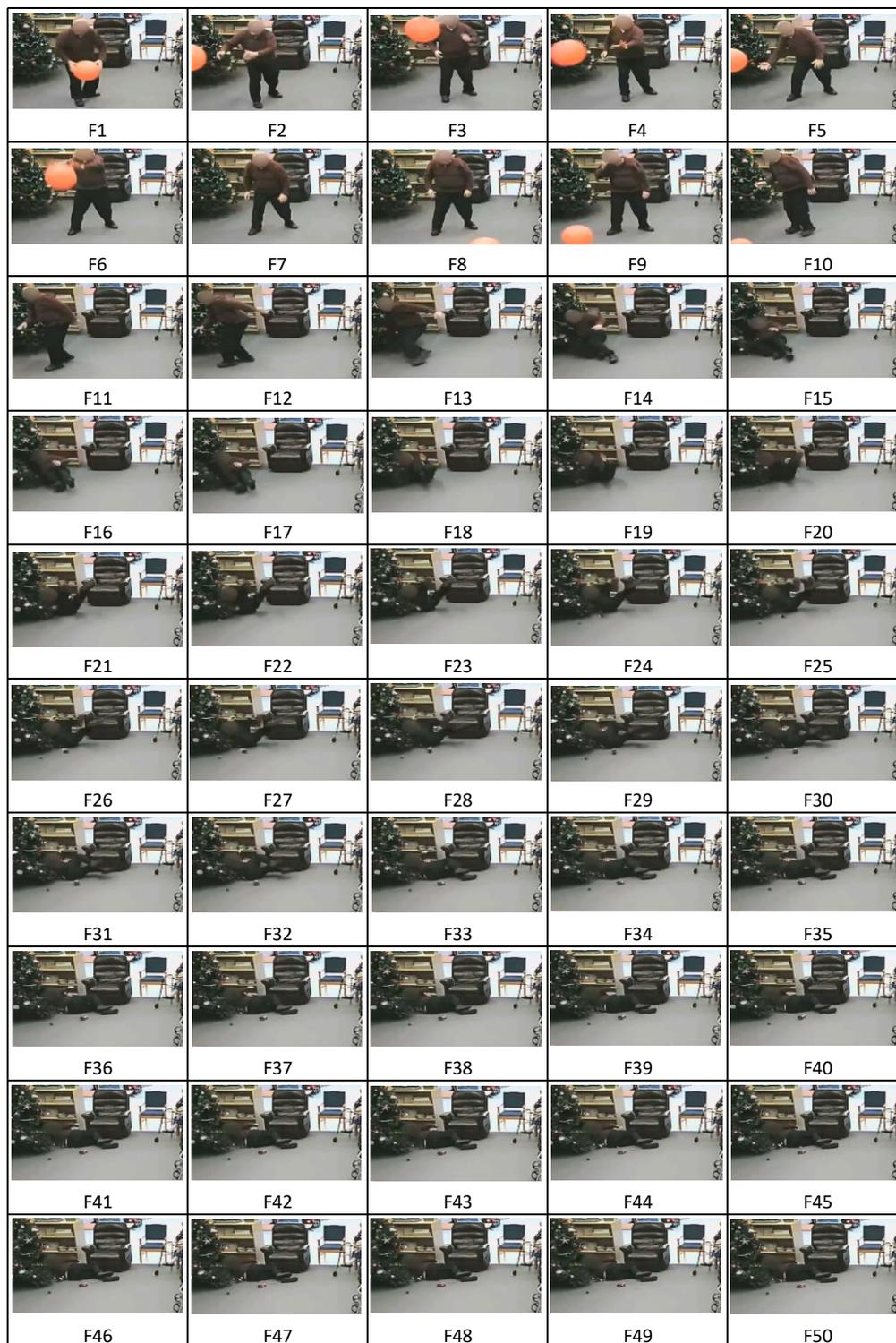


Figura 28 - *Frames* da queda de um idoso durante prática de exercício.

Fonte: Autor

<sup>3</sup> <https://www.youtube.com/watch?v=3mDmkOxprN0>

As silhuetas de deslocamento do corpo humano resultantes de cada subtração *frame a frame*, e os respectivos ângulos de inclinação  $\theta$  das elipses aproximadas, estão ilustradas na Figura 29.

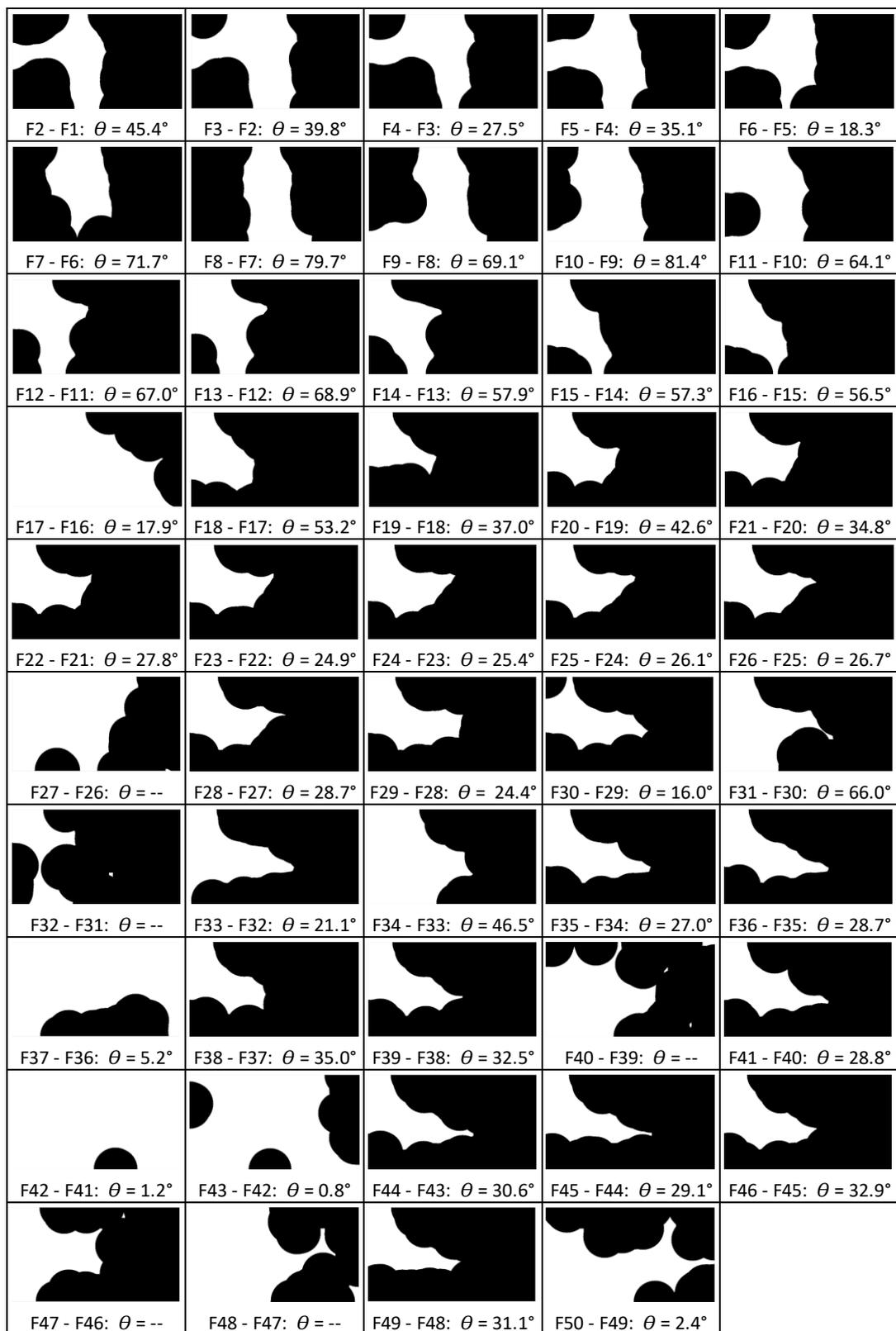


Figura 29 - Silhuetas das subtrações dos *frames* da Figura 28 e suas respectivas inclinações.

Fonte: Autor.

A Tabela 7 exibe os caracteres obtidos para cada silhueta da Figura 29 de acordo com a função de conversão expressa na equação (11).

Tabela 7 – Caracteres obtidos da aplicação 02.

SEQUÊNCIA DE CARACTERES OBTIDOS				
1: $f(45,4^\circ) = b$	2: $f(39,8^\circ) = b$	3: $f(27,5^\circ) = c$	4: $f(35,1^\circ) = b$	5: $f(18,3^\circ) = c$
6: $f(71,7^\circ) = a$	7: $f(79,7^\circ) = a$	8: $f(69,1^\circ) = a$	9: $f(81,4^\circ) = a$	10: $f(64,1^\circ) = a$
11: $f(67,0^\circ) = a$	12: $f(68,9^\circ) = a$	13: $f(57,9^\circ) = a$	14: $f(57,3^\circ) = a$	15: $f(56,5^\circ) = a$
16: $f(17,9^\circ) = c$	17: $f(53,2^\circ) = a$	18: $f(37,0^\circ) = b$	19: $f(42,6^\circ) = b$	20: $f(34,8^\circ) = c$
21: $f(27,8^\circ) = c$	22: $f(24,9^\circ) = c$	23: $f(25,4^\circ) = c$	24: $f(26,1^\circ) = c$	25: $f(26,7^\circ) = c$
26: -- = c	27: $f(28,7^\circ) = c$	28: $f(24,4^\circ) = c$	29: $f(16,0^\circ) = c$	30: $f(66,0^\circ) = a$
31: -- = a	32: $f(21,1^\circ) = c$	33: $f(46,5^\circ) = b$	34: $f(27,0^\circ) = c$	35: $f(28,7^\circ) = c$
36: $f(5,2^\circ) = c$	37: $f(35,0^\circ) = c$	38: $f(32,5^\circ) = c$	39: -- = c	40: $f(28,8^\circ) = c$
41: $f(1,2^\circ) = c$	42: $f(0,8^\circ) = c$	43: $f(30,6^\circ) = c$	44: $f(29,1^\circ) = c$	45: $f(32,9^\circ) = c$
46: -- = c	47: -- = c	48: $f(31,1^\circ) = c$	49: $f(2,4^\circ) = c$	

Devido ao fato de alguns ângulos de inclinação serem indeterminados, as silhuetas das posições 26, 31, 39, 46 e 47 da tabela, foram atribuídos aos caracteres da posição anterior.

A quantidade de silhuetas obtidas em cada zona de posição para esta aplicação está ilustrada na Figura 30. As zonas *a* e *b* possuem 13 e 6 silhuetas respectivamente, já a zona *c* é composta por 30 frames por corresponder a maior duração do vídeo, que representa o idoso caído.

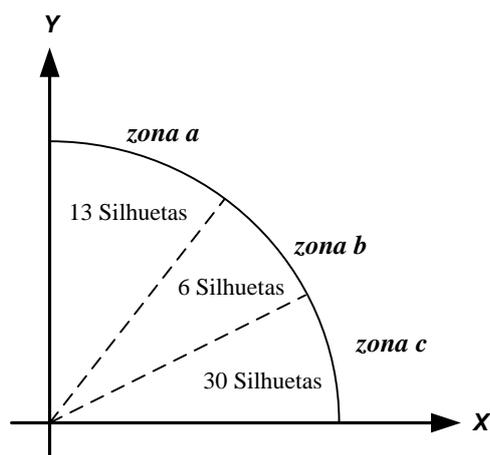


Figura 30 – Distribuição das Silhuetas por zonas durante a aplicação 02.

Fonte: Autor.



A sequência de *frames* da queda do idoso esta ilustrada na Figura 32, enumerados na sequência de  $F_1$  a  $F_{50}$ .

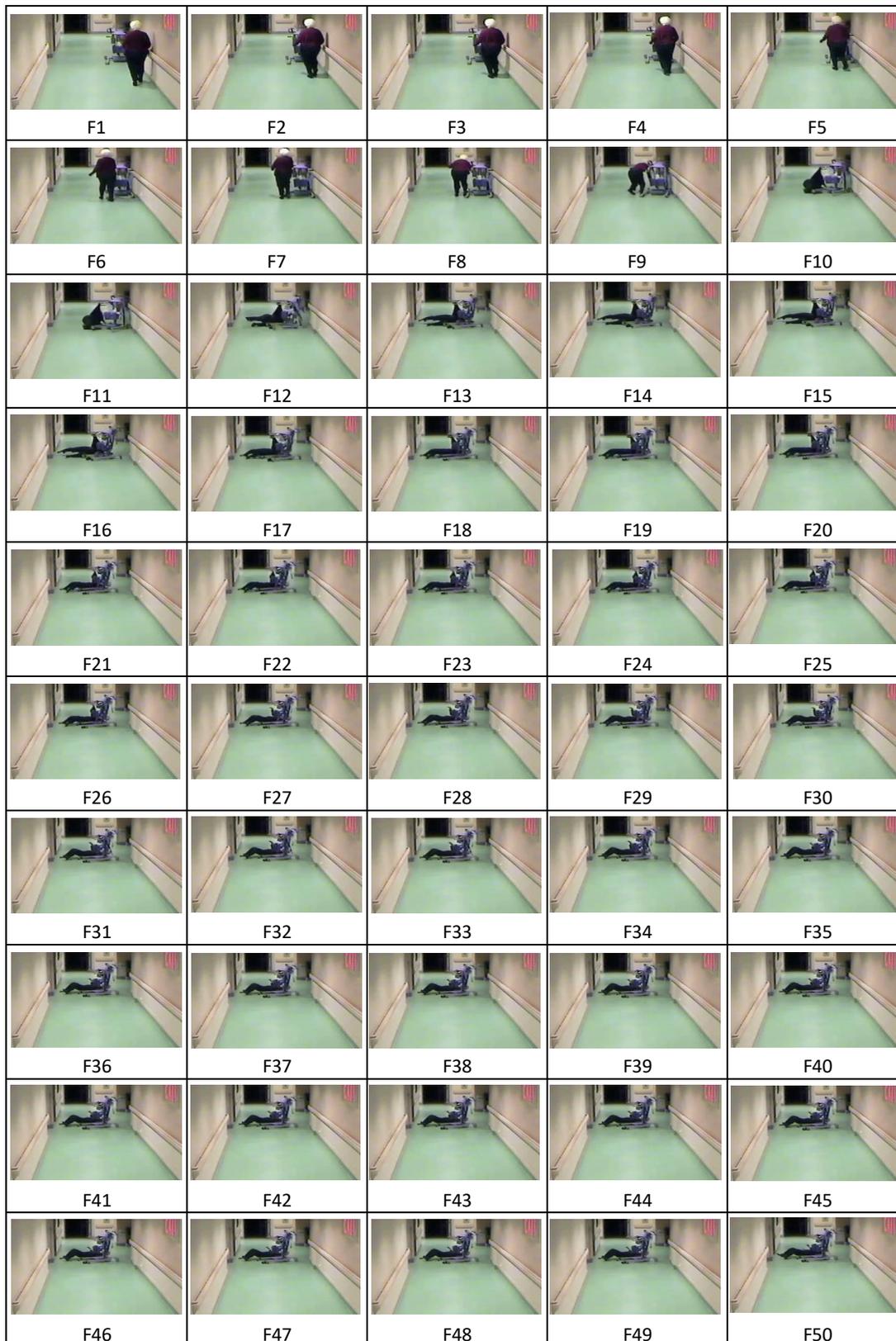


Figura 32 - *Frames* da queda de um idoso após tropeçar em um carrinho de limpeza.

Fonte: Autor.

As silhuetas de deslocamento do corpo humano resultante de cada subtração frame a frame, e os respectivos ângulos de inclinação  $\theta$  das elipses aproximadas, estão ilustradas na Figura 33.

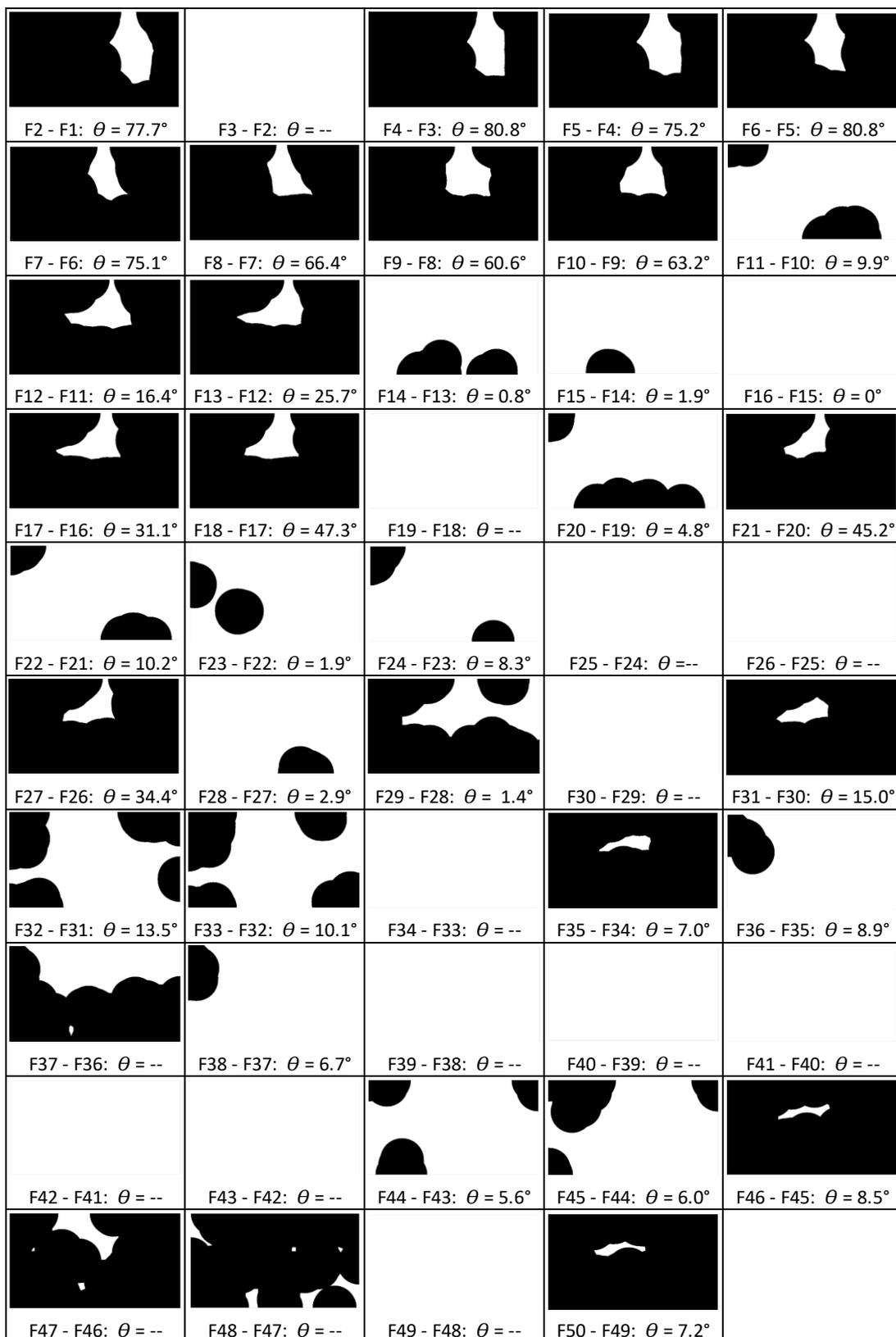


Figura 33 - Silhuetas das subtrações dos *frames* da Figura 32 e suas respectivas inclinações.  
Fonte: Autor.

A Tabela 9 exhibe os caracteres obtidos para cada silhueta da Figura 33.

Tabela 9 – Caracteres obtidos da aplicação 03.

SEQUÊNCIA DE CARACTERES OBTIDOS				
1: $f(77,7^\circ) = a$	2: $-- = a$	3: $f(80,8^\circ) = a$	4: $f(75,2^\circ) = a$	5: $f(80,8^\circ) = a$
6: $f(75,1^\circ) = a$	7: $f(66,4^\circ) = a$	8: $f(60,6^\circ) = a$	9: $f(63,2^\circ) = a$	10: $f(9,9^\circ) = c$
11: $f(16,4^\circ) = c$	12: $f(25,7^\circ) = c$	13: $f(0,8^\circ) = c$	14: $f(1,9^\circ) = c$	15: $-- = c$
16: $f(31,1^\circ) = c$	17: $f(47,3^\circ) = c$	18: $-- = c$	19: $f(4,8^\circ) = c$	20: $f(45,2^\circ) = b$
21: $f(10,2^\circ) = c$	22: $f(1,9^\circ) = c$	23: $f(8,3^\circ) = c$	24: $f(0^\circ) = c$	25: $-- = c$
26: $f(34,4^\circ) = c$	27: $f(2,9^\circ) = c$	28: $f(1,4^\circ) = c$	29: $f(0^\circ) = c$	30: $f(15,0^\circ) = c$
31: $f(13,5^\circ) = c$	32: $f(10,1^\circ) = c$	33: $-- = c$	34: $f(7,0^\circ) = c$	35: $f(8,9^\circ) = c$
36: $-- = c$	37: $f(6,7^\circ) = c$	38: $-- = c$	39: $-- = c$	40: $-- = c$
41: $-- = c$	42: $-- = c$	43: $f(5,6^\circ) = c$	44: $f(6,0^\circ) = c$	45: $f(8,5^\circ) = c$
46: $-- = c$	47: $-- = c$	48: $-- = c$	49: $f(7,2^\circ) = c$	

As silhuetas das posições 2, 15, 18, 25, 33, 36, 38, 39, 40, 41, 42, 46, 47 e 48, tiveram seus ângulos indeterminados e receberam os mesmos caracteres da posição anterior.

A quantidade de silhuetas pertencentes a cada zona de posição está ilustrada na Figura 34. As zonas *a* e *b* são compostas por 9 e 1 silhuetas respectivamente, já a zona *c* é composta por 39 silhuetas.

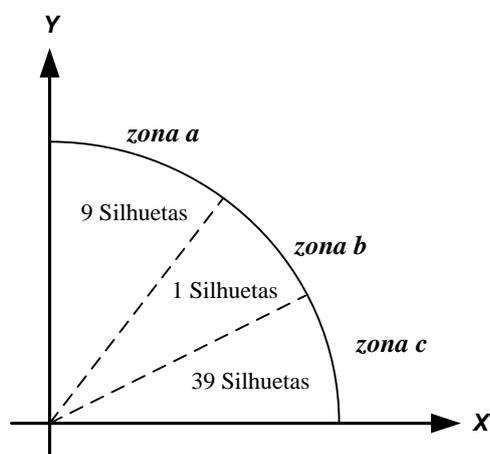


Figura 34 – Distribuição das Silhuetas por zonas durante a aplicação 03.

Fonte: Autor.

A Figura 35 ilustra o gráfico das posições versus os caracteres obtidos, nele pode ser visto que os nove primeiros caracteres correspondem ao movimento “andando”, a partir do vigésimo-primeiro caractere foram constituídas palavras referentes à “queda” a qual foi reconhecida.

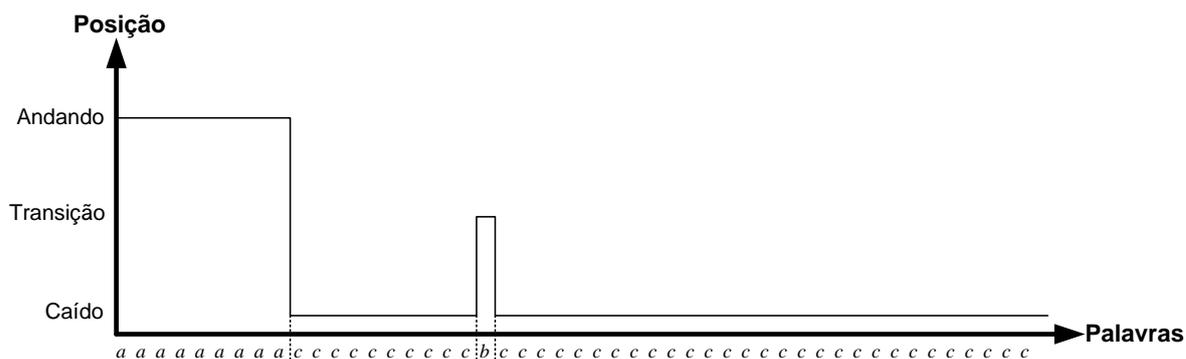


Figura 35 - Posições/movimentos obtidos pelo autômato para a aplicação 03.

Fonte: Autor

A matriz de confusão referente à aplicação 03 é mostrada na Tabela 10, onde observa-se que 39 silhuetas foram classificadas corretamente e 10 silhuetas tiveram classificação errada. A *acurácia* para esta aplicação foi de 86,4%.

Tabela 10. Matriz de confusão referente aos resultados da aplicação 03.

Silhuetas da aplicação 03		OBTIDAS NA APLICAÇÃO		
		Andando	Transição	Caído
CLASSIFICAÇÃO REAL	Andando	9	0	10
	Transição	0	1	0
	Caído	0	0	29

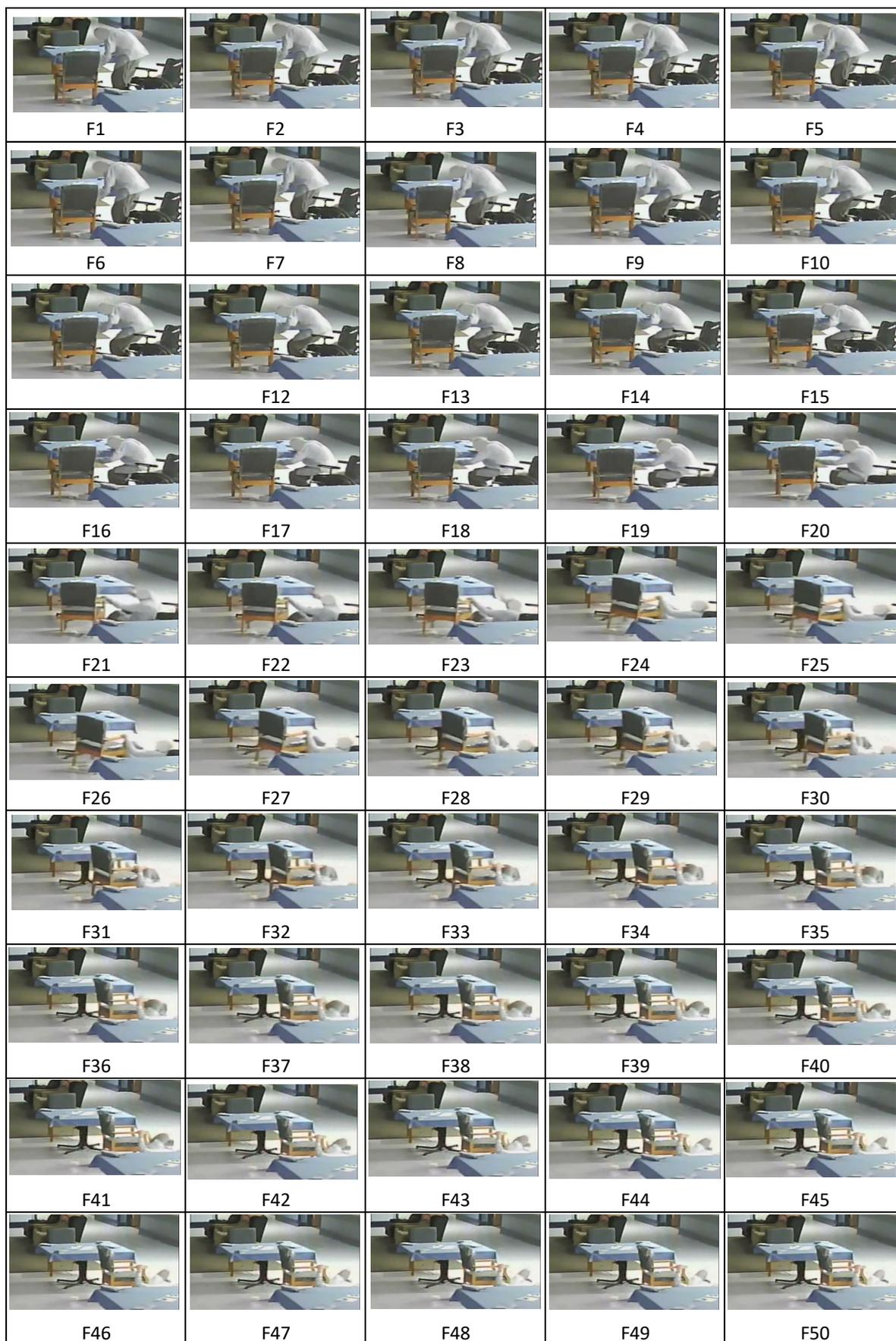
O principal fator que influenciou na classificação errada das silhuetas foi o movimento conjunto do carrinho de limpeza durante a queda do idoso, porém não afetou no reconhecimento da queda pelo método.

#### 4.2.5 Aplicação 04: Idoso cai ao tentar levantar da cadeira

Esta aplicação foi realizada durante o monitoramento de uma residência e flagra o momento da queda de um idoso que estava sentado na mesa e tenta mudar para a cadeira de rodas. Quando o idoso entra em contato com cadeira de rodas, a mesma rola para trás e não consegue fornecer suporte necessário para completar a transferência. O vídeo pode ser obtido *on-line*<sup>5</sup>.

A sequência de *frames* que representam o monitoramento está ilustrada na Figura 36, enumerados na sequência de  $F_1$  a  $F_{50}$ .

<sup>5</sup> [https://www.youtube.com/watch?v=Od\\_RgAP8ojk](https://www.youtube.com/watch?v=Od_RgAP8ojk)



**Figura 36 - Idoso cai ao levantar da cadeira.**  
**Fonte: Autor.**

As silhuetas de deslocamento do idoso e seus respectivos ângulos de inclinação  $\theta$  estão ilustrados na Figura 37.

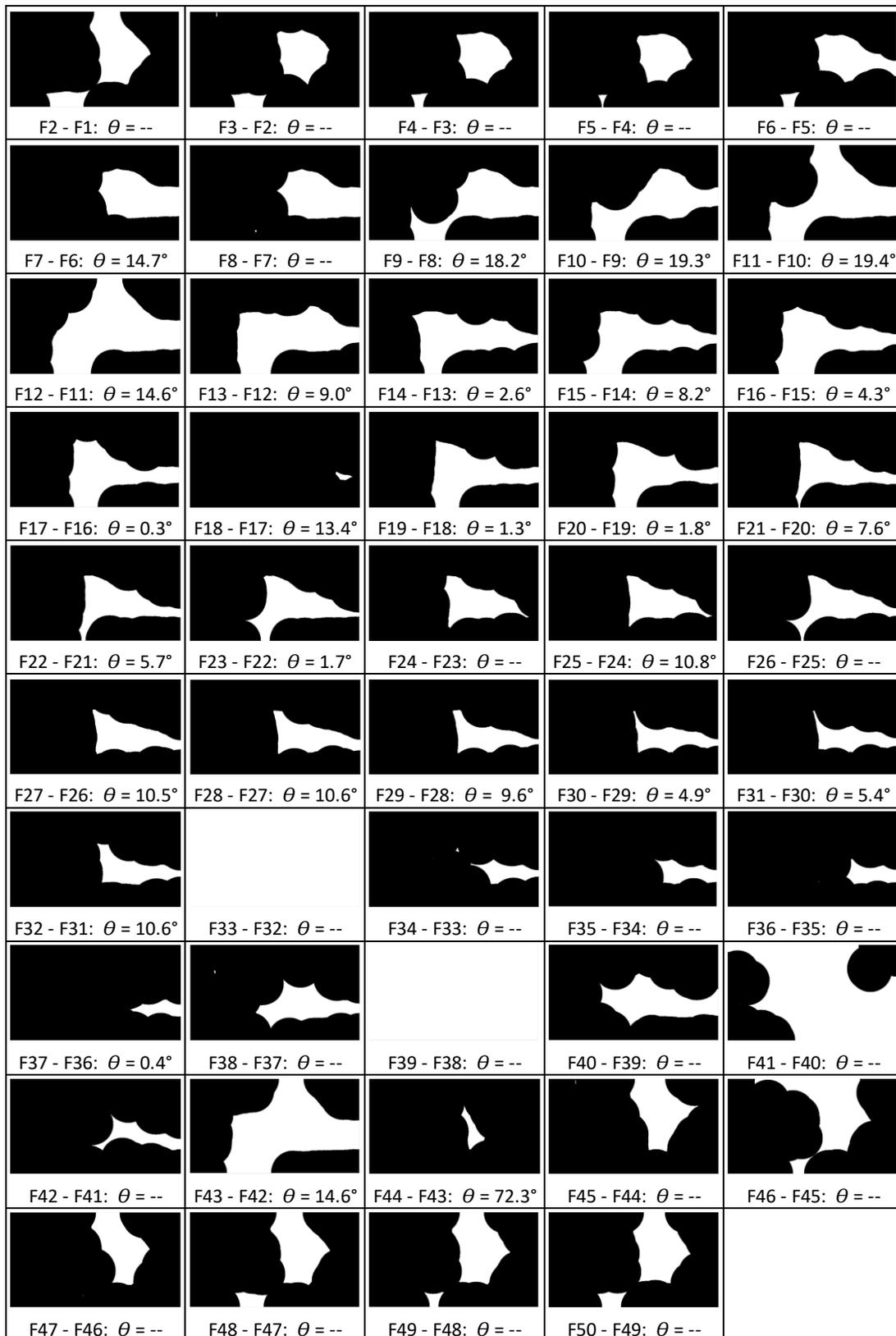


Figura 37 - Silhuetas obtidas da subtração dos *frames* da Figura 36.  
Fonte: Autor.

A Tabela 11 exibe os caracteres obtidos para cada silhueta da Figura 37.

Tabela 11 – Caracteres obtidos da aplicação 04.

SEQUÊNCIA DE CARACTERES OBTIDOS				
1: -- = c	2: -- = c	3: -- = c	4: -- = c	5: -- = c
6: $f(14,7^\circ) = a$	7: -- = a	8: $f(18,2^\circ) = a$	9: $f(19,3^\circ) = a$	10: $f(19,4^\circ) = a$
11: $f(14,6^\circ) = a$	12: $f(9,0^\circ) = a$	13: $f(2,6^\circ) = a$	14: $f(8,2^\circ) = a$	15: $f(4,3^\circ) = a$
16: $f(0,3^\circ) = a$	17: $f(13,4^\circ) = a$	18: $f(1,3^\circ) = a$	19: $f(1,8^\circ) = a$	20: $f(7,6^\circ) = a$
21: $f(5,7^\circ) = a$	22: $f(1,7^\circ) = a$	23: -- = a	24: $f(10,8^\circ) = a$	25: -- = a
26: $f(10,5^\circ) = a$	27: $f(10,6^\circ) = a$	28: $f(9,6^\circ) = a$	29: $f(4,9^\circ) = a$	30: $f(5,4^\circ) = a$
31: $f(10,6^\circ) = a$	32: $f(0^\circ) = a$	33: -- = a	34: -- = a	35: -- = a
36: $f(0,4^\circ) = a$	37: -- = a	38: $f(0^\circ) = a$	39: -- = a	40: -- = a
41: -- = a	42: $f(14,6^\circ) = a$	43: $f(72,3^\circ) = c$	44: -- = c	45: -- = c
46: -- = c	47: -- = c	48: -- = c	49: -- = c	

Os ângulos das silhuetas 1, 2, 3, 4, 5, 7, 23, 25, 33, 34, 35, 37, 39, 40, 41, 44, 45, 46, 47, 48 e 49 foram indeterminados, a eles foram atribuídos os mesmos caracteres da posição anterior.

A quantidade de silhuetas obtidas em cada zona de posição está ilustrada na Figura 38. As zonas *a* e *b* são compostas por 13 e 0 silhuetas respectivamente, a maior quantidade de silhuetas encontra-se majoritariamente na zona *c* composta por 36 silhuetas.

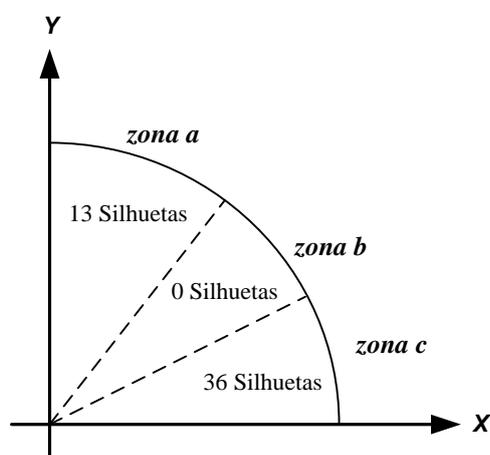


Figura 38 – Distribuição das silhuetas por zonas durante a aplicação 04.

Fonte: Autor.

O gráfico da sequência de caracteres obtidos versus as posições é ilustrado na Figura 39, nela pode ser visto que os cinco primeiros caracteres correspondem ao movimento *andando*, logo após, são gerados caracteres correspondentes ao movimento *caído* até o quadragésimo-segundo caractere, a partir daí são gerados caracteres correspondentes ao movimento *andando*, porém não





Figura 40 - Idoso cai ao descer escadaria do ICOMP.

Fonte: Autor.

As silhuetas resultantes de cada subtração *frame a frame*, e os respectivos ângulos de inclinação  $\theta$  das elipses de contorno, estão ilustrados na Figura 41.

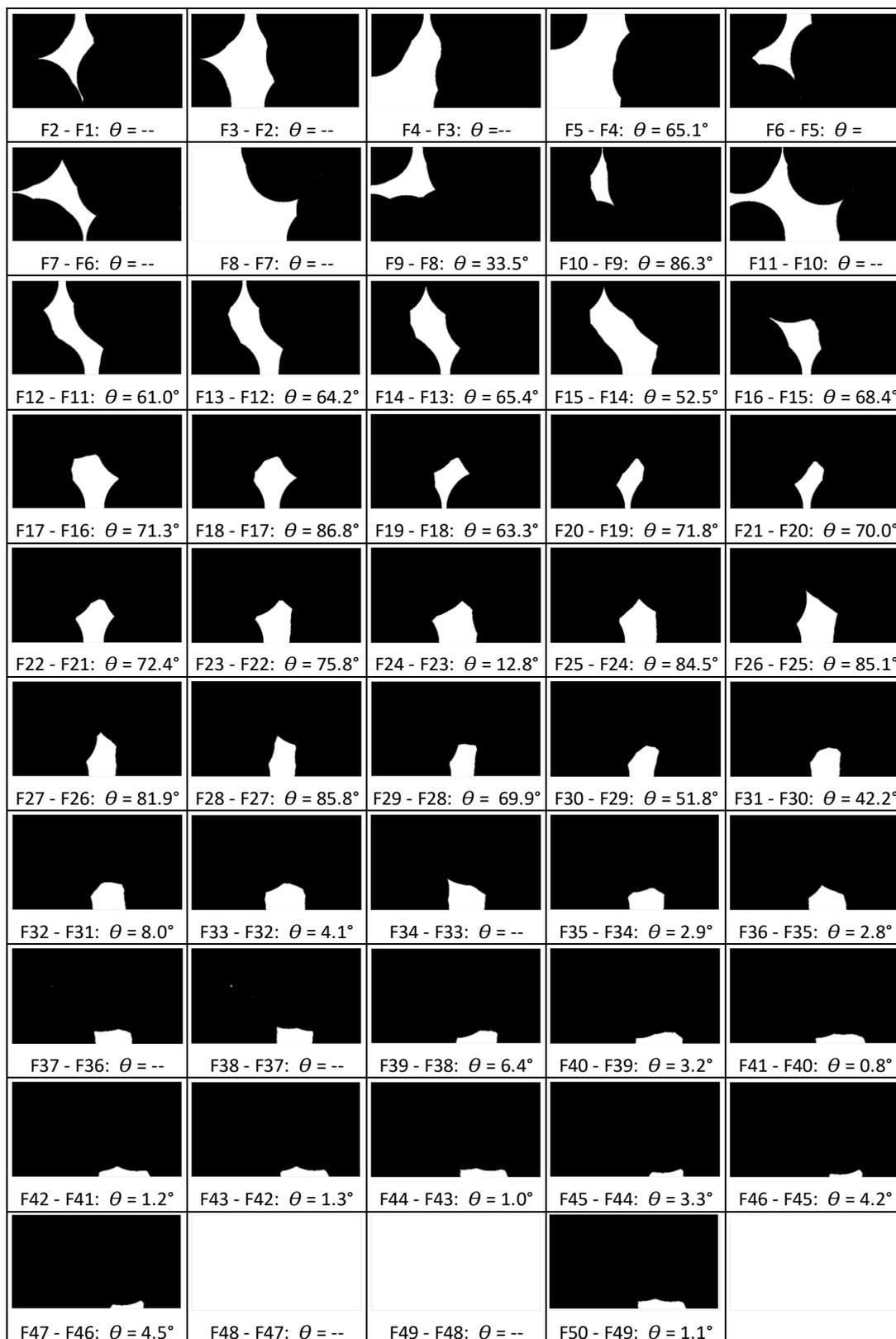


Figura 41 - Silhuetas obtidas da subtração dos *frames* da Figura 40.

Fonte: Autor.

A Tabela 13 exibe os caracteres obtidos para cada silhueta da Figura 41.

Tabela 13 – Caracteres obtidos da aplicação 05.

SEQUÊNCIA CARACTERES OBTIDOS				
1: -- = c	2: -- = c	3: -- = c	4: $f(65,1^\circ) = c$	5: -- = c
6: -- = c	7: -- = c	8: $f(33,5^\circ) = a$	9: $f(86,3^\circ) = c$	10: -- = c
11: $f(61,0^\circ) = c$	12: $f(64,2^\circ) = c$	13: $f(65,4^\circ) = c$	14: $f(52,5^\circ) = c$	15: $f(68,4^\circ) = c$
16: $f(71,3^\circ) = c$	17: $f(86,8^\circ) = c$	18: $f(63,3^\circ) = c$	19: $f(71,8^\circ) = c$	20: $f(70,0^\circ) = c$
21: $f(72,4^\circ) = c$	22: $f(75,8^\circ) = c$	23: $f(12,8^\circ) = a$	24: $f(84,5^\circ) = c$	25: $f(85,1^\circ) = c$
26: $f(81,9^\circ) = c$	27: $f(85,8^\circ) = c$	28: $f(69,9^\circ) = c$	29: $f(51,8^\circ) = c$	30: $f(42,2^\circ) = c$
31: $f(8,0^\circ) = a$	32: $f(4,1^\circ) = a$	33: -- = a	34: $f(2,9^\circ) = a$	35: $f(2,8^\circ) = a$
36: -- = a	37: -- = a	38: $f(6,4^\circ) = a$	39: $f(3,2^\circ) = a$	40: $f(0,8^\circ) = a$
41: $f(1,2^\circ) = a$	42: $f(1,3^\circ) = a$	43: $f(1,0^\circ) = a$	44: $f(3,3^\circ) = a$	45: $f(4,2^\circ) = a$
46: $f(4,5^\circ) = a$	47: $f(0,0^\circ) = a$	48: -- = a	49: $f(1,1^\circ) = a$	

Os ângulos das silhuetas 1, 2, 3, 5, 6, 7, 10, 33, 36, 37 e 48 não foram determinados, a eles foram atribuídos os caracteres da respectiva silhueta anterior.

A quantidade de silhuetas obtidas em cada zona de posição está ilustrada na Figura 42. As zonas a e b são constituídas por 28 e 0 silhuetas respectivamente, a zona c é composta por 21 silhuetas que foram suficientes para a queda ser reconhecida.

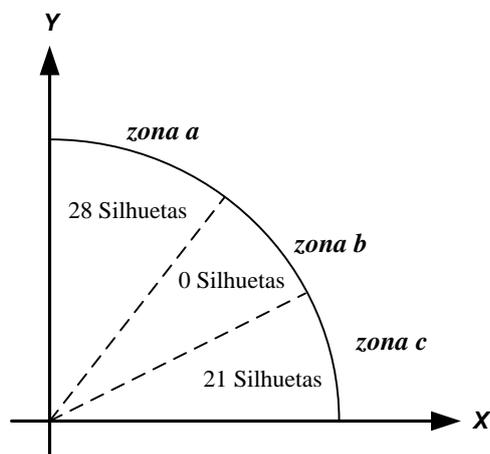


Figura 42 – Distribuição das silhuetas por zonas durante a aplicação 05.

Fonte: Autor.

O gráfico da sequência de caracteres obtidos versus as posições correspondentes é ilustrado na Figura 43, no qual pode ser visto que a partir do trigésimo-primeiro caractere são formadas palavras que representam a *queda*, que foi devidamente reconhecida pelo método.

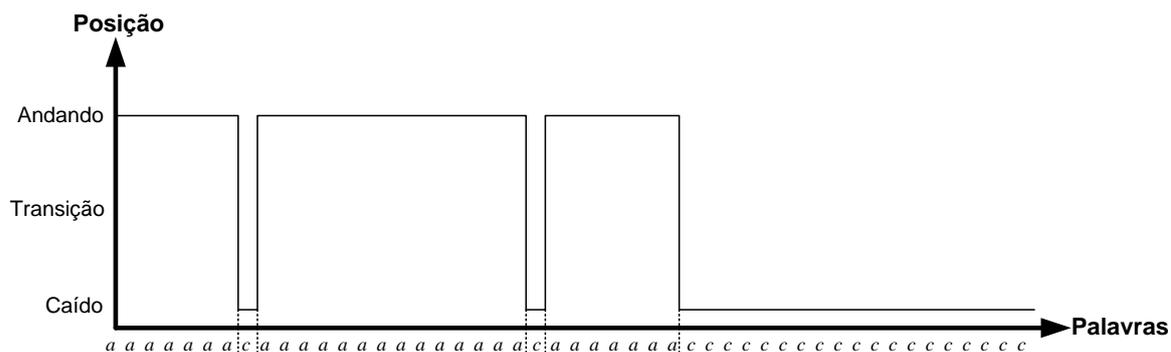


Figura 43 - Posições/movimentos obtidos pelo autômato na aplicação 05.

Fonte: Autor.

A matriz de confusão referente a esta aplicação é mostrada na Tabela 14, onde observa-se que 40 silhuetas foram classificadas corretamente e 9 silhuetas tiveram classificação errada, representando uma *acurácia* de 87,8%.

Tabela 14. Matriz de confusão referente aos resultados da aplicação 05.

Silhuetas da aplicação 05		OBTIDOS NA APLICAÇÃO		
		Andando	Transição	Caído
CLASSIFICAÇÃO REAL	Andando	21	0	1
	Transição	7	0	1
	Caído	0	0	19

O principal fator que influenciou na classificação errada das silhuetas foi à variação da intensidade do reflexo de luz emitido pelo corrimão da escadaria que é cromado e movimentou-se durante a queda do idoso gerando distorções em algumas silhuetas.

#### 4.2.7 Aplicação 06: Comparação com outros métodos

Os dados da aplicação realizada por Foroughi *et al.* (2008) foram submetidos ao método proposto neste trabalho. Os dados correspondem a 30 *frames* sintéticos de uma pessoa que anda, corre e depois sofre uma queda como ilustra a Figura 44.



**Figura 44 – Frames da aplicação realizada por Foroughi *et al.* (2008).**

**Fonte: (FOROUGHI *et al.*, 2008).**

As silhuetas obtidas na aplicação e seus respectivos ângulos de inclinação estão ilustrados na Figura 45, onde pode ser observado que as silhuetas possuem inicialmente formato na vertical, coincidindo com o deslocamento da pessoa andando, como mostram os *frames* de F<sub>1</sub> a F<sub>22</sub> da Figura 44, e a partir do *frame* F<sub>23</sub> as silhuetas correspondentes possuem um formato na horizontal correspondendo à queda sofrida pela pessoa.

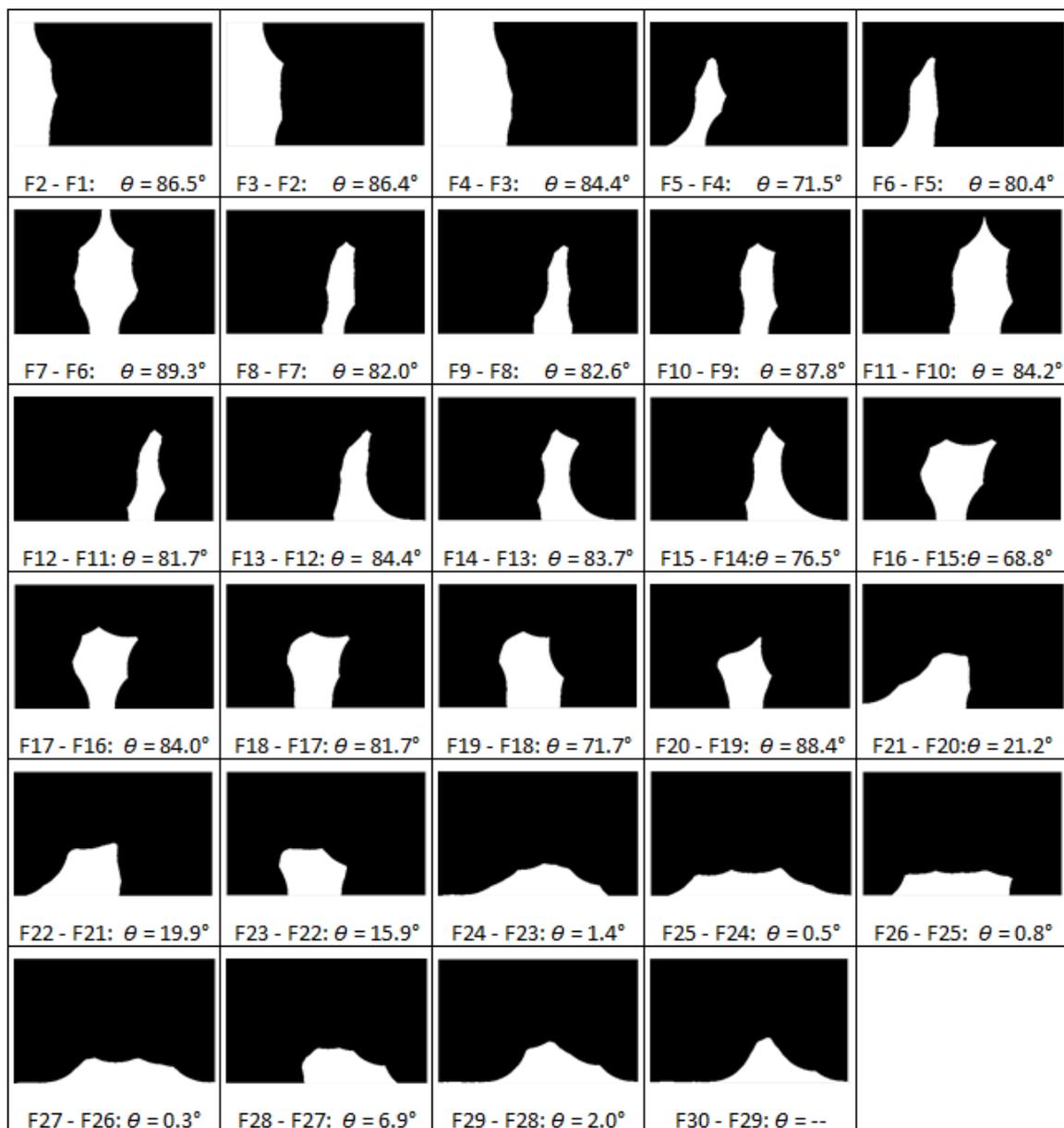


Figura 45 – Silhuetas e suas respectivas inclinações após subtração entre os *frames* da Figura 44.

Fonte: Autor.

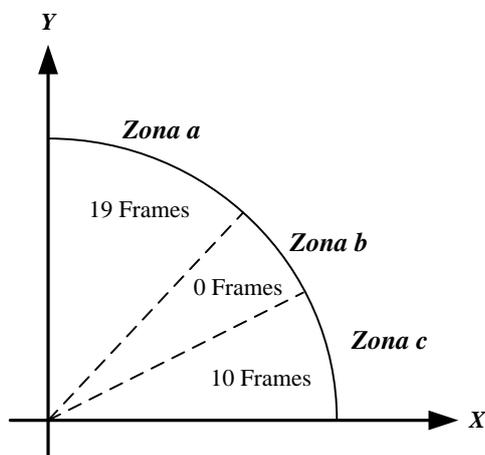
A Tabela 15 mostra a seqüência de caracteres obtidos pela função de conversão aplicada a cada silhueta da Figura 45.

Tabela 15 – Caracteres obtidos da aplicação 06.

SEQÜÊNCIA DE CARACTERES OBTIDOS				
1: $f(86.5^\circ) = a$	2: $f(86.4^\circ) = a$	3: $f(84.4^\circ) = a$	4: $f(71.5^\circ) = a$	5: $f(80.4^\circ) = a$
6: $f(89.3^\circ) = a$	7: $f(82.0^\circ) = a$	8: $f(82.6^\circ) = a$	9: $f(87.8^\circ) = a$	10: $f(84.2^\circ) = a$
11: $f(81.7^\circ) = a$	12: $f(84.4^\circ) = a$	13: $f(83.7^\circ) = a$	14: $f(76.5^\circ) = a$	15: $f(68.8^\circ) = a$
16: $f(84.0^\circ) = a$	17: $f(81.7^\circ) = a$	18: $f(71.7^\circ) = a$	19: $f(88.4^\circ) = a$	20: $f(21.2^\circ) = c$
21: $f(19.9^\circ) = c$	22: $f(15.9^\circ) = c$	23: $f(1.4^\circ) = c$	24: $f(0.5^\circ) = c$	25: $f(0.8^\circ) = c$
26: $f(0.3^\circ) = c$	27: $f(6.9^\circ) = c$	28: $f(2.0^\circ) = c$	29: $f(-- ) = c$	

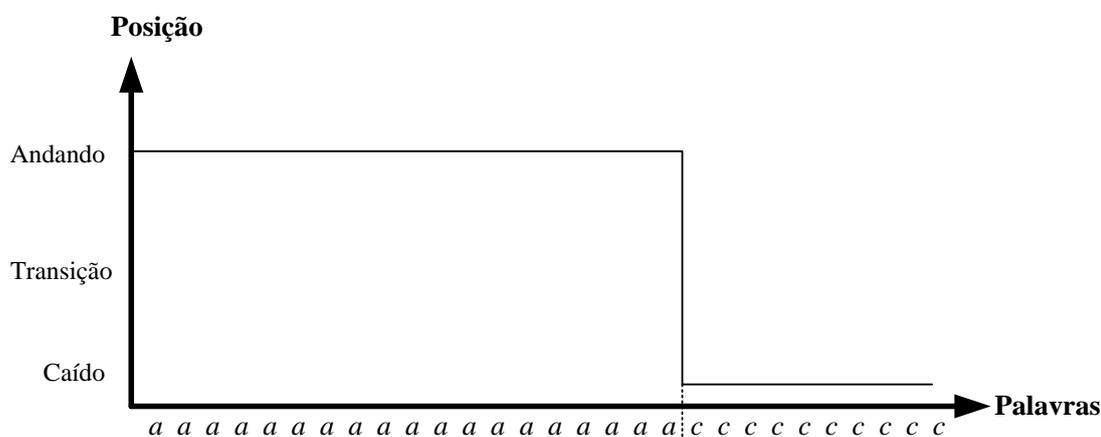
A última silhueta da Figura 45 recebeu o mesmo caractere da silhueta anterior pelo fato de ser constituída por mais de uma região de formação, como pode ser constatado na Tabela 15.

A Figura 46 ilustra a distribuição de silhuetas por zonas, a *zona b* não teve nenhuma silhueta, ou seja, o método não conseguiu detectar a transição de andando para a queda.



**Figura 46 - Distribuição das silhuetas por zonas durante a aplicação 06.**  
Fonte: Autor.

A Figura 47 ilustra as posições obtidas de acordo com a sequência de caracteres gerados.



**Figura 47 - Posições/movimentos obtidos pelo autômato na aplicação 06.**  
Fonte: Autor.

Por tratar-se de dados sintéticos, o método classificou apenas 3 *frames* erroneamente, os quais correspondem a transição de andando para caído como mostrado na Tabela 16, no entanto a queda foi reconhecida.

Tabela 16. Matriz de confusão referente aos resultados da aplicação 06.

Silhuetas da aplicação 06		OBTIDOS NA APLICAÇÃO		
		Andando	Transição	Caído
CLASSIFICAÇÃO REAL	Andando	18	0	0
	Transição	1	0	2
	Caído	0	0	8

A métrica utilizada por Foroughi *et al.* (2008) foi a *acurácia*, os autores obtiveram uma *acurácia* de 91,12% no reconhecimento dos *Frames* da Figura 45. Já o método proposto neste trabalho obteve uma *acurácia* 93,1% para a mesma aplicação e também reconheceu a queda, como mostra a comparação na Tabela 17.

Tabela 17 – Resultado da comparação entre os métodos.

Comparação		
Autor	<i>acurácia</i>	Reconheceu a queda?
(FOROUGHY <i>et al.</i> , 2008)	91,12%	Sim
Método proposto	93,10%	Sim

O método proposto neste trabalho obteve *acurácia* superior ao método de Foroughi *et al.* (2008) para esta aplicação, comprovando a confiabilidade do método.

### 4.3 ANÁLISE DOS EXPERIMENTOS

Durante os experimentos, observou-se que as quedas caracterizam-se pela interação do idoso com algum objeto da residência, como cadeira, suporte de apoio, árvore de natal, entre outros, que conseqüentemente também movimentam-se durante a queda gerando muitas silhuetas com mais de uma região de fechamento, que impossibilita a determinação de seu ângulo de inclinação. Então foi necessário ajustar o valor do raio de abrangência  $R$  do elemento estruturante. O ajuste foi de  $R = 30$  para  $R = 80$  nas *operações morfológicas*, que concentrou as várias silhuetas em apenas uma região de fechamento, possibilitando a determinação de seu ângulo de inclinação.

A matriz de confusão referente à classificação das silhuetas de todas as aplicações é mostrada na Tabela 18, onde observa-se que 216 silhuetas (diagonal principal) foram classificadas corretamente e 62 silhuetas tiveram classificação errada. O principal fator que influenciou na

classificação errada das silhuetas foi o movimento conjunto de objetos, pessoas e animais, durante a queda do idoso, isso devido o fato de tratar-se de aplicações em cenas reais de monitoramento.

**Tabela 18 – Matriz de confusão referente ao total de silhuetas.**

Total de Silhuetas		OBTIDOS NA APLICAÇÃO		
		Andando	Transição	Caído
CLASSIFICAÇÃO REAL	Andando	69	3	33
	Transição	8	13	8
	Caído	9	1	134

A posição que teve melhor classificação foi à *queda* em seguida foi *andando* e em terceiro foi à *transição*, isto porque a zona de *transição* está entre as zonas *caído* e *andando* e daí o movimento de transição assemelhar-se as outras posições. A classificação é melhor expressa de acordo com as métricas estabelecidas, que terão seus resultados exibidos posteriormente.

Para serem obtidas as métricas de avaliação do método, primeiramente é obtido o resultado da classificação da matriz de confusão mostrado na Tabela 19.

**Tabela 19 - Resultado de classificação.**

Quantidade de silhuetas	Verdadeiro Positivo (VP)	Falso Negativo (FN)	Falso Positivo (FP)	Verdadeiro Negativo (VN)
Andando	75	37	16	166
Transição	13	13	4	264
Caído	146	10	40	98
<b>TOTAL</b>	234	60	60	528

O resultado da Tabela 19 é obtido de acordo com as regras da matriz de classificação definida na Tabela 20.

**Tabela 20 – Regras da matriz de classificação.**

Quantidade		OBTIDOS PELO MÉTODO PROPOSTO	
		VERDADEIRO	FALSO
CLASSIFICAÇÃO REAL	VERDADEIRO	Verdadeiro Positivo (VP)	Falso Negativo (FN)
	FALSO	Falso Positivo (FP)	Verdadeiro Negativo (VN)

Fonte: (REVISTABW, 2017).

A partir do resultado da classificação, as métricas para avaliação do método (definidas na Seção 4.1.1) são calculadas, como mostra a Tabela 21 e discutidas posteriormente.

**Tabela 21 - Métricas para avaliação do método e valores obtidos.**

<b>MÉTRICAS</b>	<b>VALOR</b>
<i>acurácia</i>	0,85
<i>sensibilidade</i>	0,78
<i>especificidade</i>	0,89
<i>eficiência</i>	0,83
<i>Valor Preditivo Positivo-VPP</i>	0,78
<i>Valor Preditivo Negativo-VPN</i>	0,89
<i>coeficiente de correlação (<math>\phi</math>)</i>	0,67

A *acurácia* do método de 0,85 reflete um resultado pertinente para predições corretas do método. Dentre todas as aplicações realizadas, em apenas uma (aplicação 04) o método não conseguiu reconhecer a queda do idoso, isto devido ao movimento simultâneo de objetos durante o evento. A *acurácia* do método pode ser melhorada se alguns ajustes forem realizados, como a identificação de regiões verdadeiras de movimento que pode ser caracterizada por um tamanho mínimo de área de movimento.

O valor de 0,78 para a *sensibilidade* do método reflete a uma capacidade de indicar corretamente o monitoramento do movimento. Esse valor é aceitável e condiz com a *acurácia* destacando a robustez da abordagem desenvolvida. Esse valor também pode ser melhorado por meio da caracterização mais precisa do movimento individual.

A capacidade de predição do método é expressa pelo valor de 0,89 para a métrica de *especificidade*. A abordagem foi próxima do padrão verdadeiro (1) ao identificar corretamente os movimentos andando, desequilibrando e caído. Não se atendo a outras condições de movimento mais específicas. A *eficiência* de 0,83 condiz com os valores de *acurácia* e da *sensibilidade* e reflete a boa capacidade de decisão do método.

O *VPP* (Valor Preditivo positivo) e o *VPN* (Valor Preditivo Negativo) indicam a capacidade do método de encontrar valores positivos e negativos. O valor do *VPP* de 0,78 indica uma boa capacidade do método em obter valores verdadeiros positivos. O valor do *VPN* de 0,89 indica que o método possui bom desempenho em classificar os valores negativos corretamente.

O *phi*, o coeficiente de correlação de Matthews, indica um bom desempenho geral para o método para classificação. Esta estatística, assim como a *eficiência*, tenta resumir a qualidade da tabela de contingência em um único valor numérico passível de ser comparado. Esse coeficiente varia de -1 (predição inversa) a 1 (predição perfeita), passando pelo 0 (predição aleatória média). Logo, um valor de 0,67, está bem acima do limite médio e mais próximo do valor ideal. Esse valor destaca a qualidade dos resultados em aplicações que foram realizadas para cenas reais em que ocorreram movimentos simultâneos de outras pessoas, animais e objetos. Enfatiza-se que melhores resultados serão obtidos por meio do isolamento do movimento do alvo (o idoso) em relação a outros movimentos na cena.

---

## CONCLUSÕES E TRABALHOS FUTUROS

### 5.1 CONCLUSÕES

Neste trabalho desenvolveu-se um método baseado em análise sintática de dados visuais para o monitoramento e reconhecimento de situações de risco, em especial a ocorrência de quedas, de pessoas idosas. Criou-se um modelo de reconhecimento de padrões visuais que utiliza gramáticas formais obtidas a partir da identificação de primitivas formadas por padrões visuais do movimento.

O uso de gramáticas formais mostrou-se conveniente para a descrição dos movimentos que caracterizam a queda de pessoas idosas. Os dados visuais representam primitivas que são convertidas em símbolos (caracteres), que por sua vez formam as palavras que são analisadas sintaticamente.

As aplicações experimentais comprovaram a aplicabilidade e viabilidade do método. Foram realizadas cinco aplicações experimentais obtidas a partir de vídeos reais de monitoramento que apresentam queda de pessoas idosas em ambiente interno e uma aplicação comparando com outro método, sendo que o método proposto nesta pesquisa obteve melhor *acurácia*. O método apresentou satisfatórios níveis de *Acurácia*, *Sensibilidade*, *Especificidade*, *Eficiência*, Valor Preditivo Negativo, Valor Preditivo Positivo e coeficiente de Correlação para reconhecimento da queda e movimentos do idoso.

Com base nesses resultados o método destaca as seguintes vantagens sobre as abordagens existentes na literatura:

- Generalização da descrição do movimento. A partir da definição das palavras, pode-se inferir o reconhecimento de qualquer movimento que possa ser descrito;

- Não requer uso de dispositivos como os acelerômetros que necessitam de uma série de cuidados e ocasionam desconforto, visto que alguns necessitam estarem fixos ao corpo do idoso e que suas baterias sejam recarregadas constantemente;
- Não apresenta restrições que podem inviabilizar a aplicação de forma generalizada, por exemplo, a pessoa deve estar próxima da câmera;
- Não utiliza câmeras especiais como o *kinect*. Algumas abordagens reconhecem a queda apenas se o movimento ocorrer em determinada direção em relação à posição da câmera;
- Reconhece “quedas suaves”. A maioria das abordagens que utilizam dados visuais reconhecem a queda apenas se o movimento for brusco, ou seja, possivelmente não reconhecem a queda de um idoso que esteja passando mal e apoie-se em alguma coisa antes de cair lentamente;
- O custo de implementação é relativamente baixo;
- Reconhece a queda caso alguma coisa mova-se além da pessoa caindo. Na maioria das quedas o idoso interage com alguma coisa, como por exemplo tropeçar em algum objeto;
- Não utiliza classificadores estatísticos ou métodos que podem exigir um elevado poder computacional como as redes neurais.

O método apresentou as seguintes limitações:

- Reconhece apenas uma pessoa;
- Oclusões diminuem o nível de *acurácia*;
- Inviável para ambientes com baixa luminosidade.
- Inviável para ambientes externos.

## 5.2 SUGESTÕES PARA TRABALHOS FUTUROS

Como sugestões para trabalhos futuros considera-se explorar outros parâmetros da elipse envoltória da silhueta humana como: taxa de variação e localização do centroide, taxa de variação da área, inclinação, comprimento do eixo e semi-eixo, excentricidade, perímetro, extremos, entre outros parâmetros dos quais podem ser extraídas características para inclusão de outras sentenças

gramaticais que favoreçam a expansão do método para reconhecimento de outras situações de risco do idoso como: subir em algum banco ou objeto de apoio, realização de alguma tarefa perigosa, estar em algum lugar considerado perigoso, como próximo a uma escada ou piscina, no espaço reservada para o cão que pode derrubá-lo, reconhecer a presença de algum objeto ou fator perigoso, por exemplo um carrinho de limpeza deixado no corredor, entre outras atividades de risco.

Sugere-se ainda a ampliação do método para reconhecimento simultâneo de múltiplas atividades entre pessoas e integração do método em dispositivos de comunicação e em redes de monitoramento providas em ambientes de computação ubíqua. Considera-se ainda a necessidade de fornecer maior *acurácia* ao método por meio de abordagens estocásticas.

### 5.3 CONSIDERAÇÕES FINAIS

A Organização Mundial de Saúde (OMS) declarou que nas próximas décadas a população mundial com mais de 60 anos vai passar dos atuais 841 milhões para 2 bilhões até 2050, tornando as doenças crônicas e o bem-estar da terceira idade novos desafios de saúde pública global. Assim, é prioritário que os sistemas de saúde devam encontrar estratégias eficazes para resolver os problemas enfrentados pela população mundial mais envelhecida, evitando a perda de qualidade de vida.

O elevado número de idosos no mundo atual tem favorecido o desenvolvimento de sistemas para monitoramento e atenção à saúde dessas pessoas. São necessárias estratégias para melhorar a prevenção e o gerenciamento de condições crônicas, disponibilizando cuidados de excelência acessíveis a todos os idosos, levando em consideração o ambiente físico e social.

Este trabalho forneceu uma contribuição inicial no contexto da pesquisa científica na área de Visão Computacional do PPGE da Universidade Federal do Amazonas. É certo que a efetivação futura da integração da tecnologia aqui iniciada irá aprimorar consideravelmente os resultados e o uso dos sistemas de monitoramento de idosos em situação de risco. Porém, é mais evidente ainda que este trabalho se agrupe a uma linha de atuação que influenciará decisivamente o desenvolvimento de novas soluções para a melhoria da qualidade de vida das pessoas idosas.

- 
- AGGARWAL, J. K. e RYOO, M. S. Human Activity Analysis: A Review. *ACM Computing Surveys*, v. 43, n. 3, pp.1–43, 2011.
- AKANSHA, U.; SHAIENDRA, M.; SINGH, N. Analytical review on video-based human activity recognition. 3rd International Conference on Computing for Sustainable Global Development (INDIACom). *IEEE Conference Publications*, pp. 3839 – 3844, 2016.
- ALVES, J. E. D. *et al.* Estrutura etária, bônus demográfico e população economicamente ativa no Brasil: cenários de longo prazo e suas implicações para o mercado de trabalho. *Textos para Discussão CEPAL-IPEA*, 2010. Brasília, 2010.
- ALWAN, M. *et al.* A smart and passive floor-vibration based fall detector for elderly. *2nd Information and Communication Technologies*, v. 1, pp. 1003–1007, 2006.
- AMOR, B. B.; SU, J. e SRIVASTANA, A. Action recognition using rate-invariant analysis of skeletal shape trajectories. *IEEE Transactions on Pattern Analysis and machine Intelligence*, v. 38, n. 1, pp. 1–13, 2016.
- BOURKE, A. e LYONS, G. A threshold-based fall-detection algorithm using a bi-axial gyroscope sensor, *Medical Engineering & Physics*. v. 30, n. 1, pp. 84–90, 2008.
- CANDAMO, J. *et al.* Understanding Transit Scenes: A Survey on Human Behavior-Recognition Algorithms. *IEEE Transactions on Intelligent Transportation Systems*, v. 11, n. 1, 2010.
- CARVALHO, A. M. e COUTINHO, E. S. F. Demência como Fator de Risco para Fraturas Graves em Idosos. *Revista de Saúde Pública*, v. 36, n. 4. São Paulo, 2002.
- DALDOSS, M. *et al.* Activity detection using regular expressions. *Lecture Notes in Electrical Engineering 158*, pp. 91 –106. New York, 2013.
- DOLLAR, P. *et al.* Pedestrian Detection: An Evaluation of the State of the Art. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 34, n. 4, pp. 743–761, 2012.
- FOROUGH, H.; ASKI, B. S. e POURREZA, H. Intelligent Video Surveillance for Monitoring Fall Detection of Elderly in Home Environments. *Proceedings of 11th International Conference on Computer and Information Technology (ICCIT 2008)*, pp. 25-27. Bangladesh, 2008.
- GAGLIO, S.; LO, G. e MORANA, M. Human Activity Recognition Process Using 3-D Posture Data. *IEEE Transactions on Human-Machine Systems*, v. 45, n. 5, pp. 586-597. 2015.
- GANANÇA, F. F. *et al.* Circunstâncias e consequências de quedas em idosos com vestibulopatia crônica. *Revista Brasileira de Otorrinolaringologia*. v. 72, n. 3. São Paulo, 2006.

GOLZALEZ, R. C. e WOODS, R. E. *Processamento de Imagens Digitais*, Editora Edgard Blucher Ltda, 2000.

GUIMARÃES, L. H. C. T. *et al.* Comparação da propensão de quedas entre idosos que praticam atividade física e idosos sedentários. *Revista Neurociências*, v. 12, n. 2, 2004.

IBGE, Instituto Brasileiro de Geografia e Estatística. *Projeção da População do Brasil por Sexo e Idade 1980-2050*. Disponível em: <[www.seriesestatisticas.ibge.gov.br](http://www.seriesestatisticas.ibge.gov.br)>. Acesso em: Acessado em Março de 2017.

IJJINA, E. P.; MOHAN, C. K. Human Action Recognition Based on MOCAP information using convolution neural networks. *International Conference on Machine Learning and Applications (ICMLA)*, pp. 3-6. Detroit, 2014.

IVANOV, A. e BOBICK, A. F. Action recognition using probabilistic parsing. *Computer Vision and Pattern Recognition. IEEE Computer Society Conference*, pp. 196-202, 1998.

KANGAS, M. *et al.* Comparison of low-complexity fall detection algorithms for body attached accelerometers, *Gait & Posture*, v. 28, n. 2, pp. 285–291. 2008.

KARANTONIS, D. *et al.* Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring, *IEEE Transactions on Information Technology in Biomedicine*, v. 10, n. 1, pp. 156–167, 2008.

KE, S. R. *et al.* A Review on Video-Based Human Activity Recognition. *Computers*, v. 2, n. 2, p. 88–131 manuscripts. 2013.

KIM, E.; HELAL, S. e COOK, D. Human Activity Recognition and Pattern Discovery. *IEEE Pervasive Computing*. v. 9, pp. 48-53, 2010.

KO, T. A Survey on Behavior Analysis in Video Surveillance for Homeland Security Applications. In *IEEE Applied Imagery Pattern Recognition Workshop*, pp. 1–8, 2008.

LINDEN, R. e SILVA, V. N. A. L. Autômatos Finitos Não-Determinísticos para Diagnóstico de falhas em Sistemas de potência, *Revista de Sistemas de Informação da FSMA*, n. 3, p. 2-9, 2009.

LUZ, W. A. M. Análise de fatores de risco associados a queda de idosos no domicílio. *Revista portal de divulgação*, n. 31, pp. 2178-3454. São Paulo, 2013.

LENECOACHING. Disponível em <<http://www.clipartbest.com/cliparts/ncE/XrE/ncEXrEKpi.jpeg>> Acessado em Março de 2017.

MICHELETTO, T. M. G. P. O Risco do idoso pedestre nas vias urbanas. Disponível em: <<http://www.cetsp.com.br/media/96549/nt219.pdf>> Acessado em Março de 2017.

MODARRES, A. F. A. e SORYANI, M. Body posture graph: a new graph-based posture descriptor for human behavior recognition. *IET Computer Vision*. v. 7, n. 6. pp. 488-499. 2013.

MORAES, L. P. M. *et al.* Quedas de idosos no ambiente domiciliar e consequentes alterações nas atividades básicas da vida diária (avd) e nas atividades instrumentais da vida diária (aivd). *Anais IV SIMPAC*, v. 4, n. 1, p. 261-264, Viçosa, 2012.

NOURY, N. *et al.* A proposal for the classification and evaluation of fall detectors, *IRBM*, v. 29, n. 6, pp. 340–349. 2008

OGALE, A.; KARAPURKAR, A. AND ALOIMONOS, Y. View-invariant modeling and recognition of human actions using grammars. *Proceedings of the International Workshop on Dynamical Vision*, p. 115-126, 2005.

PICCARDI, M. Background subtraction techniques: a review. In *IEEE Intl. Conference on Systems, Man and Cybernetics*, v. 4, pp. 3099–3104, 2004.

PIO, J. L. S. *et al.* Reconhecimento Sintático e Estruturas formadas de padrões. De atualização em informática. Cap. 06, pp. 283 – 330. Campinas, 2004.

PIRSIAVASH, H.; RAMANAN, D. Parsing videos of actions with segmental grammars. *IEEE Conference on Computer Vision and Pattern Recognition*. pp. 1063-6919, 2014.

POPPE, R. *Vision-Based Human Motion Analysis: An Overview*. Elsevier *Computer Vision and Image Understanding*, v. 108, pp. 4–18, 2007.

REVISTABW. Matriz de Confusão. *Revista Brasileira de Web: Tecnologia*. Disponível em <<http://www.revistabw.com.br/revistabw/matriz-de-confusao/>>. Acessado em Março de 2017.

ROBINOVITCH, S. N. *et al.* Video capture of the circumstances of falls in elderly people residing in long-term care: an observational study, *Eighth International Congress on Peer Review and Scientific Publication*. v. 381, pp. 47- 41, 2012.

ROBINOVITCH, S. N. Video capture of the circumstances of falls in elderly people residing in long-term care: an observational study. Disponível em: <<http://www.thelancet.com/journals/lancet/article/PIIS0140-6736%2812%2961263-X/fulltext>> Acessado em Março de 2017.

ROSANI, A.; CONCI, N. e NATALE, F. G. B. “Human behavior recognition using a context-free grammar”. *Journal of Electronic imaging*. *Journal of Electronic Imaging*, v. 23, n. 3, 2014.

ROUGIER, C.; MEUNIER, J.; ST-ARNAUD, A. e ROUSSEAU, J. Fall detection from human shape and motion history using video surveillance, *IEEE 21st International Conference on Advanced Information Networking and Applications Workshops*, v. 2, pp. 875–880, 2007.

ROUGIER, C.; ST-ARNAUD, A.; ROUSSEAU, J. e MEUNIER, J. “Video Surveillance for Fall Detection”. *Video Surveillance, InTech*, n. 20, pp. 357-382, China , 2011.

ROZESTRATEN, R. J. A. O idoso no trânsito. *Jornal da ABRAMET*. *Psicologia do trânsito*, pp. 8, São Paulo, 1993.

SANTOS, M. L. C; ANDRADE, M. C. Incidência de Quedas Relacionada aos Fatores de Riscos em Idosos Institucionalizados. *Revista Baiana de Saúde Pública*, v. 29, n.1, pp. 57-68, Bahia, 2005.

SIQUEIRA, F. V. *et al.* Prevalence of falls in elderly in Brazil: a countrywide analysis. *Caderno de Saúde Pública*, v. 27, n. 9, pp. 1819-26, Rio de Janeiro, 2011.

SIXSMITH, A e JOHNSON, N. A smart sensor to detect the falls of the elderly. *IEEE Pervasive Computing*, v. 3, n. 2, pp. 42-47, 2004.

SOILLE, P. *Morphological Image Analysis – Principles and Applications*. Springer-Verlag Berlin Heidelberg, 1999.

TANG, A. Y. C. *et al.* Fall Detection Sensor System for the Elderly. *International Journal of Advanced Computer Research*, v. 5, pp. 2277-7970, 2015.

TURAGA, P. et al. Machine Recognition of Human Activities: A Survey. *IEEE Transactions on Circuits and Systems for Video Technology*, v. 18, n. 11, pp. 1473-1488, 2008.

VIEIRA, J. N. *Introdução aos Fundamentos da Computação – Linguagens e Máquinas*. Thomson, 2006.

WEINLAND, D.; RONFARD, R. AND BOYER, E. A survey of vision-based methods for action representation, segmentation and recognition. *Elsevier Computer Vision and Image Understanding*, v. 115, n. 2, pp. 224-241, 2011.

WORRAKULPANIT, N. AND SAMANPIBOON, P. Human Fall Detection Using Standard Deviation of C-Motion Method. *Journal of Automation and Control Engineering*, v. 2, n. 4, 2014.

XIA, L.; CHIH, C. AND AGGARWAL, J. K. View Invariant Human Action Recognition Using Histograms of 3D Joints. *IEEE*, v. 978, n. 1, pp. 4673-1612, 2012.

YANG, Y.; GUHA, A.; FERMULLER, C.; ALOIMONOS, Y. AND BUILDING, W. A Cognitive System for Understanding Human Manipulation Actions. *Advances in Cognitive Systems*, v. 3, pp. 67-86, 2014.

YAO, B. AND ZHU, S. C. Learning deformable action templates from cluttered videos. In *IEEE Intl. Conference on Computer Vision*, pp. 1507-1514, 2009.