



**UNIVERSIDADE ESTADUAL DO CEARÁ**  
**CENTRO DE CIÊNCIAS E TECNOLOGIA**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO**  
**MESTRADO ACADÊMICO EM CIÊNCIA DA COMPUTAÇÃO**

**ANTÔNIO JOSE FERNANDES ANDRADE**

**ESTUDO COMPARATIVO DE DESEMPENHO ENTRE CASAMENTO DE  
*TEMPLATE*, FILTRO DE KALMAN E FILTRO DE PARTÍCULAS PARA O  
RASTREAMENTO VISUAL DO OBJETO**

**FORTALEZA – CEARÁ**

**2015**

ANTÔNIO JOSE FERNANDES ANDRADE

ESTUDO COMPARATIVO DE DESEMPENHO ENTRE CASAMENTO DE *TEMPLATE*,  
FILTRO DE KALMAN E FILTRO DE PARTÍCULAS PARA O RASTREAMENTO VISUAL  
DO OBJETO

Dissertação apresentada ao Curso de Mestrado Acadêmico em Ciência da Computação do Programa de Pós-Graduação em Ciência da Computação do Centro de Ciências e Tecnologia da Universidade Estadual do Ceará, como requisito parcial à obtenção do título de mestre em Ciência da Computação. Área de Concentração: Ciência da Computação

Orientador: José Everardo Bessa Maia

FORTALEZA – CEARÁ

2015

Dados Internacionais de Catalogação na Publicação

Universidade Estadual do Ceará

Sistema de Bibliotecas

Andrade, Antônio José Fernandes.

Estudo comparativo de desempenho entre casamento de template, filtro de Kalman e filtro de partículas para o rastreamento visual do objeto [recurso eletrônico] / Antônio José Fernandes Andrade. - 2015.  
1 CD-ROM: il.; 4 ¾ pol.

CD-ROM contendo o arquivo no formato PDF do trabalho acadêmico com 76 folhas, acondicionado em caixa de DVD Slim (19 x 14 cm x 7 mm).

Dissertação (mestrado acadêmico) - Universidade Estadual do Ceará, Centro de Ciências e Tecnologia, Mestrado Acadêmico em Ciência da Computação, Fortaleza, 2015.

Área de concentração: Ciências da Computação.  
Orientação: Prof. Dr. José Everardo Bessa Maia.

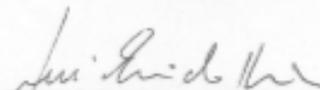
1. Rastreamento visual. 2. Casamento de Template.  
3. Filtro de Kalman. 4. Filtro de Partículas.. I.  
Título.

	<b>UNIVERSIDADE ESTADUAL DO CEARÁ - UECE</b> PRÓ-REITORIA DE PÓS-GRADUAÇÃO E PESQUISA - PROPGPq CENTRO DE CIÊNCIAS E TECNOLOGIA - CCT Mestrado Acadêmico em Ciência da Computação - MACC <small>Av. Dr. Síllas Maragábe, 1700, Campus do Baperi, Fortaleza, Ceará, Brasil. Fone: 310197266</small>	
---	--	---

**ATA DA SEPTUAGÉSIMA DEFESA PÚBLICA  
DE DISSERTAÇÃO DE Mestrado**



Aos treze dias do mês de fevereiro de dois mil e quinze, no miniauditório do prédio de Pesquisa e Pós-Graduação em Computação, do Mestrado Acadêmico em Ciência da Computação - MACC, realizou-se a sessão pública de defesa da dissertação de **ANTÔNIO JOSÉ FERNANDES ANDRADE**, aluno regularmente matriculado no Mestrado Acadêmico em Ciência da Computação - MACC, tendo como título: "Estudo Comparativo do Desempenho entre Casamento de Template, Filtro de Kalman e Filtro de Partículas para o Rastreamento Visual do Objeto". A Banca Examinadora reuniu-se no horário de 9:00h às 11:12 horas, sendo constituída pelos Professores Doutores da Universidade Estadual do Ceará: **José Everardo Bessa Maia**-(Orientador), **Thelmo Pontes de Araújo-Phd**, ambos da Universidade Estadual do Ceará-UECE, e o **Prof. Dr. Geraldo Luis Bezerra Ramalho** do Instituto Federal de Educação, Ciência e Tecnologia do Estado do Ceará-IFCE. Inicialmente o mestrando expôs seu trabalho e a seguir foi submetido à arguição pelos membros da Banca, dispondo cada membro de tempo para tal. Finalmente a Banca reuniu-se em separado e concluiu por considerar o mestrando APROVADO, por sua dissertação e sua defesa pública. Eu, Professor **Dr. José Everardo Bessa Maia**, Orientador da Dissertação e Presidente da Banca, lavro a presente Ata que será assinada por mim e demais membros. Fortaleza, 13 de fevereiro de 2015.

  
 Prof. Dr. José Everardo Bessa Maia  
 Orientador - UECE

  
 Prof. Phd. Thelmo Pontes de Araújo  
 UECE

  
 Prof. Dr. Geraldo Luis Bezerra Ramalho  
 IFCE

## **AGRADECIMENTOS**

Primeiramente a Deus, criador dos céus e da terra, que guiou meus passos e livrou-me de diversos problemas.

A minha esposa, Jany Cardoso, pelo amor, carinho e apoio.

Ao professor Dr. Everando Bessa Maia pela orientação.

A todos os meus companheiros da UECE, tanto professores, quanto colegas de turma.

E a todos que fizeram parte ativamente ou passivamente deste trabalho.

“É melhor lançar-se à luta em busca do triunfo mesmo expondo-se ao insucesso, que formar fila com os pobres de espírito, que nem gozam muito nem sofrem muito; E vivem nessa penumbra cinzenta sem conhecer nem vitória nem derrota.”

(Franklin Roosevelt)

## RESUMO

Este trabalho implementa e compara o desempenho de três abordagens para o problema do rastreamento visual automático de um objeto em sequências de vídeo. Os algoritmos são: Casamento de *Template*, Filtro de Kalman e Filtro de Partículas. Os desempenhos destes métodos são comparados entre si frente a dificuldades presentes nos vídeos: oclusão, manobra, variação de iluminação e variação na aparência do objeto. Para comparar os desempenhos são utilizadas uma medida de acurácia e uma medida de robustez. Resultados qualitativos e quantitativos são apresentados.

**Palavras-chave:** Rastreamento visual. Casamento de *Template*. Filtro de Kalman. Filtro de Partículas.

## **ABSTRACT**

This work implements and compares the performance of three approaches to the problem of automatic visual object tracking in video sequences. The algorithms are: Template Matching, Kalman Filter and Particle Filter. The performances of these methods are compared with each other face to present difficulties in the videos: occlusion maneuver, lighting variation and variation in the appearance of the object. To compare the performances an accuracy measure and a measure of robustness are used. Qualitative and quantitative results are presented.

**Keywords:** Visual tracking. Template Matching. Kalman Filter. Particulate Filter.

## LISTA DE ILUSTRAÇÕES

Figura 1 – Representação de objetos. a)Centróide, b) Multi-pontos, c) Forma retangular, d) Forma elíptica, e) Formas articuladas, f) Forma Esqueleto, g) Contorno completo do objeto, h) Ponto de controle sobre o contorno do objeto, i) Silhueta do objeto. (Figura extraída de YILMAZ et al.(2006) . . . . .	18
Figura 2 – Visão geral de um mecanismo de rastreamento. . . . .	23
Figura 3 – Exemplo de um gráfico gerado em cada quadro no rastreamento por detecção	25
Figura 4 – Determinação do objeto e a sua posição no quadro $i$ . . . . .	25
Figura 5 – Determinação da área de busca no quadro $i + 1$ . . . . .	25
Figura 6 – Diagrama do rastreamento como um problema de inferência sequencial. . .	29
Figura 7 – Etapas do Filtro de Kalman (adaptado de (BISHOP; WELCH, 2001)). . . .	33
Figura 8 – Princípio da amostragem por importância (Adaptado de (GRISSETTI; STA-CHNISS, 2007)). . . . .	35
Figura 9 – Ilustração do funcionamento do Filtro de Partículas. . . . .	37
Figura 10 – Fluxograma de rastreamento com Casamento de <i>Template</i> . . . . .	39
Figura 11 – Seleção do rosto em perfil de uma pessoa como objeto alvo . . . . .	39
Figura 12 – Fluxograma de rastreamento com Filtro de Kalman . . . . .	41
Figura 13 – Fluxograma de rastreamento com Filtro de Partículas . . . . .	43
Figura 14 – Trajetória do centro do objeto alvo no vídeo 1. . . . .	48
Figura 15 – Trajetória do centro do objeto alvo no vídeo 2. . . . .	49
Figura 16 – Trajetória do centro do objeto alvo no vídeo 3. . . . .	51
Figura 17 – Trajetória do centro do objeto alvo no vídeo 4 . . . . .	52
Figura 18 – Visualização da Acurácia. Adaptada de KRISTAN et al. (2013). . . . .	53
Figura 19 – Visualização da robustez. Adaptada de KRISTAN et al. (2013). . . . .	54
Figura 20 – Bias de Reinicialização. Adaptada de FERNANDEZ; VOJIR (2013). . . .	55
Figura 21 – Média dos valores da acurácia das quinze execuções para cada quadro do vídeo 1 . . . . .	58
Figura 22 – Imagem com os quadros do vídeo 1 para o rastreador com Casamento de <i>Template</i> . . . . .	59
Figura 23 – Imagem com os quadros do vídeo 1 para o rastreador com Filtro de Kalman	59
Figura 24 – Imagem com os quadros do vídeo 1 para o rastreador com Filtro de Partículas	59

Figura 25 – Média dos valores da acurácia das quinze execuções para cada quadro do vídeo 2 . . . . .	61
Figura 26 – Imagem com os quadros do vídeo 2 para o rastreador com Casamento de <i>Template</i> . . . . .	62
Figura 27 – Imagem com os quadros do vídeo 2 para o rastreador com Filtro de Kalman	62
Figura 28 – Imagem com os quadros do vídeo 2 para o rastreador com Filtro de Partículas	62
Figura 29 – Média dos valores da acurácia das quinze execuções para cada quadro do vídeo 3 . . . . .	64
Figura 30 – Imagem com os quadros do vídeo 3 para o rastreador com Casamento de <i>Template</i> . . . . .	65
Figura 31 – Imagem com os quadros do vídeo 3 para o rastreador com Filtro de Kalman	65
Figura 32 – Imagem com os quadros do vídeo 3 para o rastreador com Filtro de Partículas	66
Figura 33 – Média dos valores da acurácia das quinze execuções para cada quadro do vídeo 4 . . . . .	68
Figura 34 – Imagem com os quadros do vídeo 4 para o rastreador com Casamento de <i>Template</i> . . . . .	68
Figura 35 – Imagem com os quadros do vídeo 4 para o rastreador com Filtro de Kalman	69
Figura 36 – Imagem com os quadros do vídeo 4 para o rastreador com Filtro de Partículas	69
Figura 37 – Gráficos A-R de cada rastreador por vídeo. . . . .	70
Figura 38 – Resultado dos rastreadores MAPA A-R. . . . .	71

## LISTA DE TABELAS

Tabela 1 – Tempo de execução definidos pelo tamanho da área de busca . . . . .	40
Tabela 2 – Vídeos utilizados para avaliar os rastreadores . . . . .	46
Tabela 3 – Valores para as características MA, O, M e MI . . . . .	46
Tabela 4 – Rotulação do vídeo 1 – MA, O, M e MI . . . . .	47
Tabela 5 – Rotulação do vídeo 2 – MA, O, M e MI . . . . .	48
Tabela 6 – Rotulação do vídeo 3 – MA, O, M e MI . . . . .	50
Tabela 7 – Rotulação do vídeo 4 – MA, O, M e MI . . . . .	51
Tabela 8 – Resultados do vídeo 1 . . . . .	57
Tabela 9 – Teste de hipótese - vídeo 1. . . . .	57
Tabela 10 – Resultados do vídeo 2 . . . . .	60
Tabela 11 – Teste de hipótese - vídeo 2. . . . .	60
Tabela 12 – Resultados do vídeo 3 . . . . .	63
Tabela 13 – Teste de hipótese - vídeo 3. . . . .	63
Tabela 14 – Resultados do vídeo 4 . . . . .	66
Tabela 15 – Teste de hipótese - vídeo 4. . . . .	67
Tabela 16 – Média de quadros válidos dos algoritmos por vídeo . . . . .	70
Tabela 17 – Resumo dos resultados dos rastreadores. . . . .	72

## LISTA DE SÍMBOLOS

$I_i$	Imagem de um vídeo num tempo $i$ .
$N(\mu, \sigma)$	Representação de uma distribuição Normal/Gaussiana com média $\mu$ e variância $\sigma$ .
$(a, b)$	Representa as coordenadas em <i>pixels</i> de um ponto em um quadro.
$x_i$	Vetor de estado interno do objeto num tempo $i$ .
$y_i$	Vetor de medição/observação num tempo $i$ .
$p(x_i y_0, \dots, y_i)$	Representa a distribuição de probabilidade de estado a <i>posteriori</i> , obtida a partir da distribuição a <i>priori</i> com um modelo de observação $p(y_i x_i)$ .
$p(x_i y_0, \dots, y_{i-1})$	Representa a distribuição de probabilidade do estado a <i>priori</i> , obtida a partir da função de densidade de probabilidade do estado anterior $p(x_{i-1} y_0, \dots, y_{i-1})$ evoluindo em função do tempo através de $p(x_i x_{i-1})$ .
$\bar{x}_i^-$	Representa a média de $P(x_i y_0, \dots, y_{i-1})$ .
$\bar{x}_i^+$	Representa a média de $P(x_i y_0, \dots, y_i)$ .
$\sigma_i^-$	Representa o desvio padrão de $P(x_i y_0, \dots, y_{i-1})$ .
$\sigma_i^+$	Representa o desvio padrão de $P(x_i y_0, \dots, y_i)$ .
$\Sigma_i^-$	Representa a matriz de covariância de $P(x_i y_0, \dots, y_{i-1})$ .
$\Sigma_i^+$	Representa a matriz de covariância de $P(x_i y_0, \dots, y_i)$ .
$\phi_i$	Representa a acurácia de um rastreador no quadro $i$ .
$\phi(j, k)$	Representa a acurácia média de um rastreador $j$ na repetição $k$ .
$\rho A(j)$	Representa a acurácia média de um rastreador $j$ .
$\gamma(j, k)$	Representa o número de falhas de um rastreador $j$ na repetição $k$ .
$\rho R(j)$	Representa a robustez média de um rastreador $j$ .
$\tau R(j)$	Representa a robustez média absoluta de um rastreador $j$ .
$N_{rep}$	Representa a quantidade de execução do rastreador.

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	14
1.1	OBJETIVOS	14
<b>1.1.1</b>	<b>Objetivo Geral</b>	14
<b>1.1.2</b>	<b>Objetivos Específicos</b>	15
1.2	CONTRIBUIÇÕES DESTA DISSERTAÇÃO	15
1.3	ESTRUTURA DESTA DISSERTAÇÃO	15
<b>2</b>	<b>ESTRUTURAÇÃO DO PROBLEMA</b>	17
2.1	O PROBLEMA DO RASTREAMENTO EM VÍDEO	17
2.2	BREVE REVISÃO DA LITERATURA	19
2.3	CONCLUSÕES DO CAPÍTULO	21
<b>3</b>	<b>REFERENCIAL TEÓRICO</b>	22
3.1	RASTREAMENTO POR CASAMENTO DE <i>TEMPLATE</i>	24
3.2	MODELOS DINÂMICOS LINEARES DISCRETOS	26
3.3	RASTREAMENTO COMO UM PROBLEMA DE INFERÊNCIA	28
3.4	RASTREAMENTO UTILIZANDO FILTRO DE KALMAN	30
<b>3.4.1</b>	<b>O Filtro de Kalman para um vetor de estado em 1D</b>	31
<b>3.4.2</b>	<b>O Filtro de Kalman para vetor de estado geral</b>	32
3.5	RASTREAMENTO UTILIZANDO FILTRO DE PARTÍCULAS	33
<b>3.5.1</b>	<b>Sequential Importance Sampling (SIS)</b>	34
<b>3.5.2</b>	<b>Sequential Importance Resampling (SIR)</b>	36
<b>3.5.3</b>	<b>Funcionamento qualitativo do algoritmo SIR</b>	36
3.6	CONCLUSÕES DO CAPÍTULO	37
<b>4</b>	<b>IMPLEMENTAÇÃO</b>	38
4.1	SISTEMA RASTREADOR COM CASAMENTO DE <i>TEMPLATE</i>	38
4.2	SISTEMA RASTREADOR COM FILTRO DE KALMAN	41
4.3	SISTEMA RASTREADOR COM FILTRO DE PARTÍCULAS	43
4.4	CONCLUSÕES DO CAPÍTULO	45
<b>5</b>	<b>METODOLOGIA DE AVALIAÇÃO</b>	46
5.1	CARACTERIZAÇÃO DOS VÍDEOS	46
<b>5.1.1</b>	<b>Análise do Vídeo 1</b>	47

5.1.2	<b>Análise do Vídeo 2</b> . . . . .	48
5.1.3	<b>Análise do Vídeo 3</b> . . . . .	50
5.1.4	<b>Análise do Vídeo 4</b> . . . . .	51
5.2	<b>METODOLOGIA DE COMPARAÇÃO DE DESEMPENHO</b> . . . . .	53
5.2.1	<b>Acurácia</b> . . . . .	53
5.2.2	<b>Robustez</b> . . . . .	54
5.3	<b>CONCLUSÕES DO CAPÍTULO</b> . . . . .	55
6	<b>RESULTADOS E DISCUSSÃO</b> . . . . .	56
6.1	<b>RESULTADOS</b> . . . . .	56
6.1.1	<b>Vídeo 1</b> . . . . .	56
6.1.2	<b>Vídeo 2</b> . . . . .	60
6.1.3	<b>Vídeo 3</b> . . . . .	62
6.1.4	<b>Vídeo 4</b> . . . . .	66
6.2	<b>DISCUSSÃO</b> . . . . .	69
6.3	<b>CONCLUSÕES DO CAPÍTULO</b> . . . . .	72
7	<b>CONCLUSÕES E TRABALHOS FUTUROS</b> . . . . .	73
	<b>REFERÊNCIAS</b> . . . . .	74

# 1 INTRODUÇÃO

O rastreamento de objetos em vídeos aplica-se em diversas situações do dia-a-dia e é importante em áreas como o monitoramento de pessoas (crianças perdidas em um estádio ou em um shopping center) e na área de segurança (através do rastreamento de objetos ou pessoas suspeitos em aeroportos), dentre outras. Assim, ferramentas para automatizar o processo de rastreamento de objetos em vídeo possibilitando auxiliar seres humanos que trabalham nessas atividades se tornam importantes, pois na maioria das vezes são atividades repetitivas e cansativas.

O rastreamento (*tracking*) de objetos em vídeo (ou sequência de imagens) é o ramo da visão computacional que estuda o comportamento de um objeto em uma sequência de imagens, e é uma das possíveis formas de aumentar o nível de interação homem-máquina (O'HAGAN; ZELINSKY, 1997). O desenvolvimento de ferramentas que possibilitem rastrear objetos é uma tarefa complexa devido a alguns aspectos particulares que podem ocorrer, tais como: o objeto na imagem sofre alterações de escala, mudanças bruscas na trajetória do movimento, alterações na velocidade de deslocamento, oclusão, entre outras características. De forma muito geral, identificam-se três abordagens no tratamento do problema de rastreamento visual de objeto: técnicas de casamento de *template* (CT), técnicas baseadas em aprendizagem de máquina e técnicas utilizando estimação bayesiana. Este trabalho tomou casamento de *template* como método *baseline* para estudar duas propriedades dos métodos de estimação bayesiana em rastreamento visual de objeto. A primeira é o efeito de se incorporar um modelo de movimento e observação (presente em inferência bayesiana, mas não em casamento de *template*). A segunda é o efeito da hipótese de gaussianidade dos ruídos do modelo e de medição. Os algoritmos representantes das técnicas de estimação bayesiana são o Filtro de Kalman (FK) e um Filtro de Partículas (FP). Este trabalho não estuda métodos de aprendizagem de máquinas (classificação).

## 1.1 OBJETIVOS

### 1.1.1 Objetivo Geral

Implementar e comparar o desempenho dos algoritmos **Casamento de *Template***, **Filtro de Kalman** e um **Filtro de Partículas** na tarefa de rastreamento visual de objeto.

### 1.1.2 Objetivos Específicos

- a) Implementar os algoritmos Casamento de *Template*, Filtro de Kalman e Filtro de Partículas para rastreamento de um objeto em sequências de imagens.
- b) Analisar e comparar os resultados obtidos pelos algoritmos Casamento de *Template*, Filtro de Kalman e Filtro de Partículas, quando o objeto rastreado está submetido a oclusão, mudança de aparência, mudança de iluminação e manobra, de tal forma a identificar para quais situações determinada técnica de rastreamento é mais adequada.

## 1.2 CONTRIBUIÇÕES DESTA DISSERTAÇÃO

As principais contribuições desta dissertação para a comunidade foram:

- Desenvolvimento de programas em MATLAB aplicados para o rastreamento visual de objeto em vídeo. Estes programas serão publicados para a comunidade.
- A análise, comparação e recomendação do algoritmo a ser utilizado, entre Casamento de *Template*, Filtro de Kalman e Filtro de Partículas no rastreamento do objeto, quando o mesmo está submetido a diferentes níveis de oclusão, mudança de aparência, mudança de iluminação e manobra.
- Uma análise preliminar sobre a hipótese de gaussianidade do ruído nas imagens e o efeito do modelo de movimento nas propriedades dos rastreadores.
- Artigo em preparação: “Object Visual tracking with occlusion and maneuver using an online Dynamic Model of the Movement” (Rastreamento Visual de Objeto com oclusão e manobra utilizando um Modelo Dinâmico online do Movimento).

## 1.3 ESTRUTURA DESTA DISSERTAÇÃO

Esta dissertação está organizada da seguinte forma:

O **capítulo 2** apresenta algumas formas de representação do objeto alvo de um rastreamento em vídeo e as dificuldades que envolve a elaboração do rastreador, tais como as de mudança de iluminação, oclusões, aparência, manobras, entre outras (seção 2.1). Na seção 2.2 são apresentados trabalhos que utilizaram técnicas de rastreamento por detecção e outros que fizeram por inferência probabilística, através do uso do Filtro de Kalman e Filtro de Partículas. Por fim, são apresentados trabalhos que fizeram a comparação entre algumas das técnicas de

rastreamento.

O **capítulo 3** mostra algumas técnicas de rastreamento de objetos em vídeo. Na seção 3.1 é apresentado o rastreamento por Casamento de *Template*. Alguns conceitos sobre modelos dinâmicos são apresentados na seção 3.2. A forma de tratar rastreamento como um problema de inferência é exposto na seção 3.3. Por fim, na seção 3.4 é apresentado o rastreamento utilizando Filtro de Kalman e na seção 3.5 o rastreamento utilizando Filtro de Partículas.

O **capítulo 4** descreve o método do rastreador por Casamento de *Template*, por Filtro de Kalman e por Filtro de Partículas. Na seção 4.1 é descrito como funciona o rastreador com Casamento de *Template*, o qual faz uso de uma biblioteca do MATLAB, a **vision.TemplateMatchir**. Na seção 4.2 é mostrado como funciona o rastreador por Filtro de Kalman, que também usa uma biblioteca chamada **vision.KalmanFilter** e o modelo dinâmico do objeto alvo para o Filtro de Kalman. Na seção 4.3 é apresentado o método do Filtro de Partículas e o modelo dinâmico de cada partícula.

O **capítulo 5** descreve a metodologia utilizada para a comparação de desempenho dos rastreadores abordados neste trabalho. A seção 5.1 descreve a caracterização dos vídeos e na seção 5.2 é apresentada a medida de acurácia e a medida de robustez.

O **capítulo 6** apresenta os resultados do rastreamento com as técnicas de Casamento de *Template*, Filtro de Kalman, Filtro de Partículas e análise comparativa, quantitativa e qualitativa dos resultados.

Por fim, no **capítulo 7** contém as conclusões deste trabalho, além de apontar os trabalhos futuros.

## 2 ESTRUTURAÇÃO DO PROBLEMA

### 2.1 O PROBLEMA DO RASTREAMENTO EM VÍDEO

O rastreamento visual em uma sequência de imagens é a tarefa que consiste em seguir a posição de um alvo de interesse (objeto alvo), determinando sua trajetória bidimensional, no plano de imagem, utilizando informações advindas de sensores óticos, além de informações sobre os ruídos e de todo o conhecimento *a priori* disponível sobre a dinâmica dos objetos alvos desejados (SÄRKKÄ et al., 2004).

Alguns pontos, para serem decididos, são comuns na elaboração de um algoritmo de rastreamento visual de objetos (YILMAZ et al., 2006), são eles: a **representação do objeto**, os **descritores** que são utilizados na representação e o tipo de mecanismo de **detecção**.

#### Representação do Objeto

O objeto alvo pode ser representado por diferentes formas, que permite uma relação com as suas características geométricas. Segundo YILMAZ et al. (2006), as principais formas de representações são:

- a) **Pontos:** O objeto é representado por um ponto, centróide, ou conjunto de pontos, de acordo com a Figura 1 (a) e (b), respectivamente.
- b) **Formas geométricas primitivas:** Utilizam quadrados, círculos, retângulos, elipses, etc., para a representação de objetos. A Figura 1 (c) e (d) apresenta exemplos.
- c) **Formas articuladas:** O objeto é representado por várias partes unidas entre si por meio de juntas, conforme a Figura 1 (e).
- d) **Esqueleto:** Objetos compostos por um eixo medial, extraído de sua silhueta, conforme Figura 1 (f) (BALLARD; BROWN, 1992)
- e) **Contorno e Silhueta:** O contorno define os limites em torno do objeto e a silhueta é a região interna do contorno, conforme a Figura 1 (g),(h) e (i) respectivamente.

#### Descritores

Os descritores são características extraídas de uma imagem, possibilitando, através da seleção adequada das características, separar o objeto alvo no espaço de solução. Existem diversas características que podem ser extraídas de uma imagem, tais como:

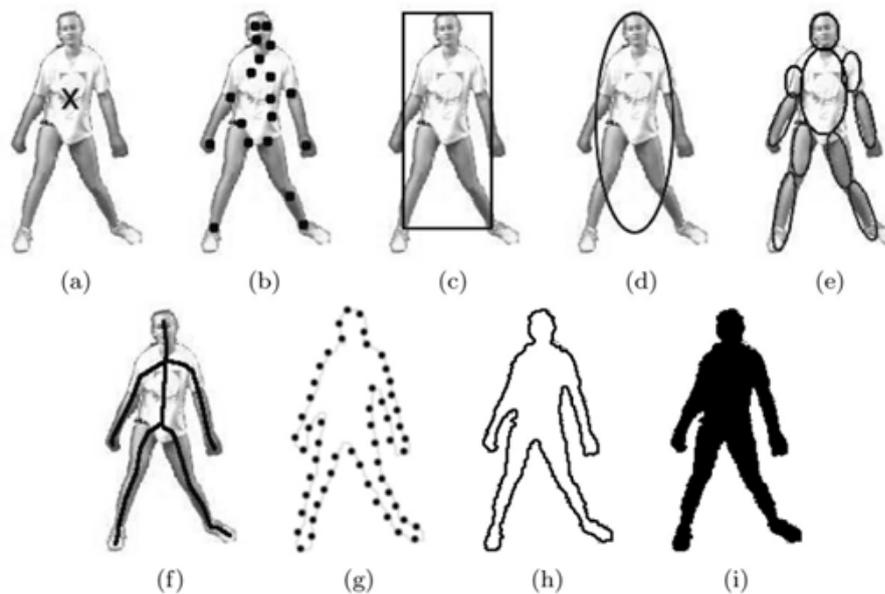


Figura 1 – Representação de objetos. a) Centróide, b) Multi-pontos, c) Forma retangular, d) Forma elíptica, e) Formas articuladas, f) Forma Esqueleto, g) Contorno completo do objeto, h) Ponto de controle sobre o contorno do objeto, i) Silhueta do objeto. (Figura adaptada de YILMAZ et al. (2006).

- **Cor:** a representação de um objeto pode ser em diferentes espaços de cor (RGB, HSV, YCbCr, YUV). A utilização da cor como característica é relativamente independente do tamanho, orientação e resolução da imagem.
- **Textura:** São as diferenças locais nos níveis de intensidade. GONZALES; WOODS (2002) citam a existência de três principais formas de descrever textura: espectral, estrutural e estatísticas.

## Detecção

No rastreamento visual de objetos é necessário um mecanismo de detecção, que pode usar informações de diversos quadros ou apenas de um único quadro. Alguns dos mecanismos de detecção são:

**Subtração de Plano de Fundo:** Um método de subtração de plano de fundo simples consiste em utilizar um único quadro que não contenha objetos de interesse para representar a cena em fundo (*background*) e, posteriormente, subtrair do quadro que contém os objetos alvos o quadro selecionado como background. Algumas limitações desse método são: as câmeras devem ser estacionárias e não deve ter variação de iluminação.

**Segmentação:** Esse método consiste em dividir uma imagem em várias regiões que possuem a mesma similaridade. O objetivo desse método é fazer com que os objetos alvos em uma imagem tenham os seus *pixels* agrupados e destacados dos demais.

**Aprendizagem Supervisionada:** Essa técnica detecta o objeto alvo através de um modelo matemático gerado por um algoritmo de aprendizagem de máquina, onde é utilizado um conjunto de exemplos do objeto para treinar o modelo.

**Reconhecimento de Padrões:** Essa técnica consiste em encontrar um padrão ou objeto alvo (uma parte específica de uma imagem) em uma imagem maior. A técnica por reconhecimento de padrões geralmente é mais rápida, pois não necessita processar a imagem inteira, apenas as características da imagem.

Algumas dificuldades devem ser contornadas na elaboração de um rastreador, pois o mesmo deve ser suficientemente genérico e, normalmente, possuir um desempenho compatível com aplicações de tempo real. Alguns dos desafios na elaboração do rastreador são:

- Manobras do objeto alvo.
- Ruídos nas imagens.
- Mudanças na aparência do objeto alvo.
- A forma complexa do objeto alvo e suas articulações naturais.
- Oclusão parcial ou total.
- Movimento da câmera.
- Mudança de iluminação.
- Mudanças de escala do objeto alvo.
- Necessidade de processamento em tempo real.

## 2.2 BREVE REVISÃO DA LITERATURA

Na literatura existem diferentes técnicas para realizar o rastreamento de objetos em vídeo, desde aquelas que funcionam através da detecção visual das características nos *pixels* de cada quadro do vídeo até as que abordam o modelo dinâmico do movimento do objeto a ser rastreado. Esta subseção faz uma breve revisão de alguns trabalhos da área com o objetivo de mostrar uma visão da grande variedade de métodos investigados na solução do problema de rastreamento de objetos em vídeo.

As técnicas que utilizam o modelo dinâmico do movimento do objeto são abordagens probabilísticas baseadas em um conjunto de informações analisadas em cada quadro, sendo

as abordagens que apresentam os melhores resultados (NORVIG; RUSSELL, 2004). Devido às informações possuírem ruídos e incertezas, pois o objeto alvo sofre alterações de escala, mudanças de aparência, oclusões e mudança de iluminação, entre outras características, este problema se ajusta bem para ser modelado em um processo de estimação. O processo de estimação é a tarefa de prever o estado de um sistema estocástico utilizando informações que são funções do estado e das observações (GREWAL; ANDREWS, 2011).

No rastreamento por detecção, o objeto alvo é representado por uma região no plano da imagem, extraído informações dos *pixels*, componentes da região, como pode ser visto no trabalho de WIDYAWAN; NUGROHO (2012), que implementa um sistema de detecção baseado no casamento de *template* para rastrear objetos, sendo adaptado para as mudanças de iluminação e movimento da câmera.

Por outro lado, existem diversos trabalhos com rastreadores que utilizam modelo dinâmico do movimento, através de abordagens probabilísticas para rastrear objetos, como foi feito em AMARAL; TOZZI (2013) que implementa um sistema de rastreamento de pessoas e tratamento de oclusões por detecção de cabeças. O sistema utiliza o Filtro de Kalman para realizar o rastreamento e a implementação é dividido em duas etapas principais: detecção do *foreground* e determinação das trajetórias. Na primeira etapa é realizada a subtração do fundo, remoção de sombras e operações morfológicas e são detectadas as regiões correspondentes às pessoas a serem seguidas. Na segunda etapa são usadas métricas de associação para associar as pessoas seguidas às pessoas detectadas na imagem.

BEM (2007) emprega dois Filtros de Kalman para rastrear diferentes objetos (tais como figuras geométricas, pessoas e robôs móveis), um para estimação do movimento e outro para estimação do contorno dos objetos em uma sequência de imagens, apresentando os resultados obtidos em testes com sequências de imagens controladas e reais, garantindo um bom desempenho e eficácia e é livre de modelos em relação às características do objeto alvo. No entanto, possui uma complexidade maior na definição das variáveis de estado do sistema.

Alguns autores fazem comparação entre o Filtro de Kalman e o Filtro de Partículas. Por exemplo, ARULAMPALAM et al. (2002) apresenta a comparação de diversas variantes do Filtro de Partículas com o Filtro de Kalman Estendido. No entanto, esta comparação não é realizada com aplicações de rastreamento em vídeo. ILHA (2009) implementa o Filtro de Kalman e o Filtro de Partícula para rastrear uma bola de futebol, fazendo uma comparação do desempenho obtido por cada um deles, levando em consideração a linearidade de movimento,

mudanças de direção e velocidade, oclusões, presença de público e distância da câmera. Sendo que esse trabalho se limita a rastrear uma bola de futebol, suas conclusões não podem ser estendidas para o caso geral.

Nesta dissertação, o desempenho do Filtro de Kalman e do Filtro de Partículas são comparados entre si e com o método do Casamento de *Template* utilizando sequências de vídeo muito gerais.

### 2.3 CONCLUSÕES DO CAPÍTULO

Este capítulo revisou importantes temas relacionados ao rastreamento de objetos em vídeo, tais como: a segmentação do objeto, as características, os métodos de detecção, os principais fatores que dificultam o rastreamento e alguns métodos para tratar o problema. Pode-se afirmar que é possível um grande número de combinações de técnicas e que ainda não há uma técnica geral aplicável a todas as situações.

### 3 REFERENCIAL TEÓRICO

Uma visão geral do mecanismo de rastreamento visual de objeto pode ser visto na Figura 2.

O objeto alvo é determinado no primeiro quadro na qual aparece. A determinação do objeto alvo pode ser baseada na seleção manual pelo o usuário, pela detecção automática do movimento ou pela detecção de um objeto específico, com as suas características já pré-definidas.

Nos quadros seguintes é utilizada alguma técnica de rastreamento com o intuito de estimar a posição do objeto alvo.

Por fim é atualizada a representação do objeto alvo. Esta etapa é necessária pelo fato de que o objeto alvo pode sofrer variações de aparência no decorrer dos quadros.

Neste capítulo serão apresentadas as técnicas de Casamento de *Template*, Filtro de Kalman e Filtro de Partículas para o problema de rastreamento de objeto em sequência de imagens. Os principais conceitos e definições foram extraídas do texto de FORSYTH; PONCE (2002).

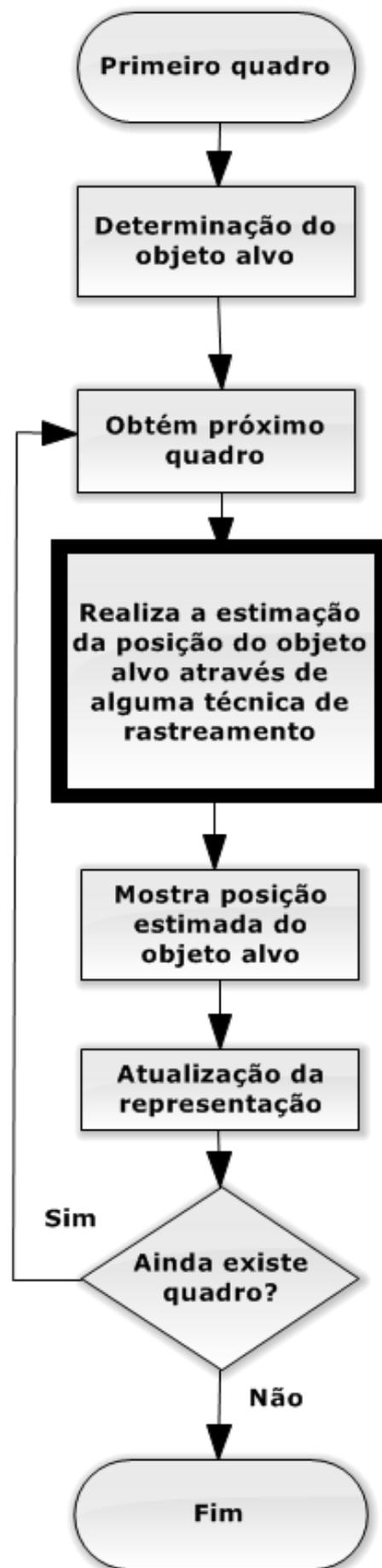


Figura 2 – Visão geral de um mecanismo de rastreamento.

### 3.1 RASTREAMENTO POR CASAMENTO DE *TEMPLATE*

Neste trabalho, um *template* é uma região de interesse na imagem. Geralmente, uma região de interesse é a representação do objeto alvo.

Após a determinação do objeto alvo e a sua representação através de uma matriz de valores, a detecção de objeto nos quadros seguintes se baseia em encontrar um conjunto de *pixels* cujas características sejam semelhantes aos *pixels* da imagem do objeto alvo. A idéia do rastreamento de objeto por Casamento de *Template*(CT), consiste em encontrar em cada um dos quadros do vídeo a matriz de valores dos *pixels* que representa a imagem do objeto alvo.

A matriz de valores mais representativa do objeto alvo pode ser encontrada por diversos métodos. Aqui é utilizada a seguinte abordagem: em um quadro  $i+1$ , é determinada uma área de busca em torno das coordenadas que o objeto alvo têm no quadro  $i$ , onde o tamanho da área de busca depende da natureza do objeto alvo. Para cada posição  $(a,b)$  dos *pixels* na área de busca é encontrado um valor de casamento, de acordo com algum critério de casamento. Alguns critérios possíveis são (FRIEMEL et al., 1995):

- Correlação entre matrizes (Corr2):

$$Corr2 = \frac{\sum_{i=1}^M \sum_{j=1}^N (A_{i,j} - \bar{A}) \cdot (B_{i,j} - \bar{B})}{\sqrt{(\sum_{i=1}^M \sum_{j=1}^N (A_{i,j} - \bar{A})^2) \cdot (\sum_{i=1}^M \sum_{j=1}^N (B_{i,j} - \bar{B})^2)}}$$

- Soma das diferenças absolutas (SAD):

$$SAD = \sum_{i=1}^M \sum_{j=1}^N |A_{i,j} - B_{i,j}|$$

- Soma dos quadrados das diferenças (SSD):

$$SSD = \sum_{i=1}^M \sum_{j=1}^N (A_{i,j} - B_{i,j})^2$$

Nestas equações,  $A$  é uma sub região da área de busca do quadro em análise,  $B$  a região do objeto alvo e  $M$  e  $N$  as suas dimensões.

Por qualquer destes critérios, este procedimento gera um gráfico como o da Figura 3.

Pode ser determinado que as coordenadas que obtiveram o maior valor de casamento passam a ser as coordenadas do objeto alvo no quadro  $i+1$ .

Resumindo, os passos do rastreamento por detecção são:

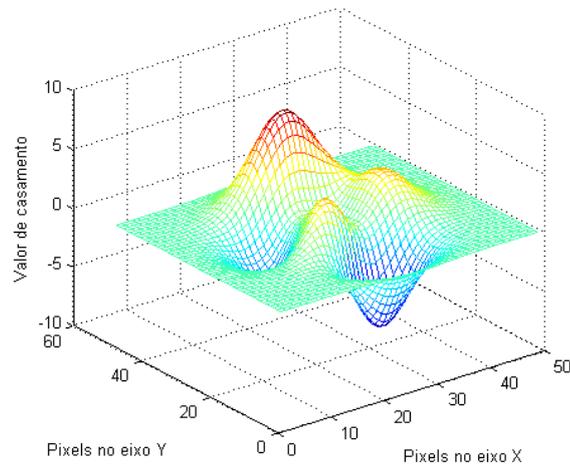


Figura 3 – Exemplo de um gráfico gerado em cada quadro no rastreamento por detecção.

1) Extraí-se duas informações do quadro  $i$ , a imagem do objeto alvo e a sua posição.  
Conforme Figura 4.

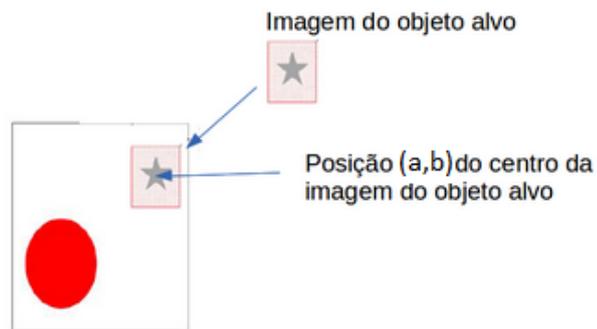


Figura 4 – Determinação do objeto e a sua posição no quadro  $i$ .

2) No quadro  $i + 1$  é determinado uma área de busca, conforme Figura 5.

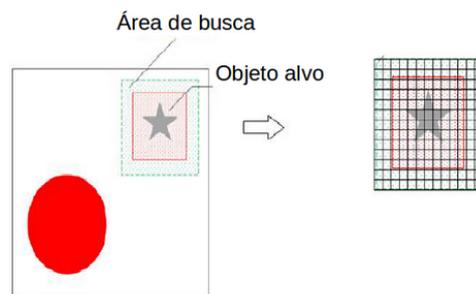


Figura 5 – Determinação da área de busca no quadro  $i + 1$ .

3) Procura-se na área de busca, através de casamento entre a matriz de valores dos *pixels* do objeto alvo com uma matriz de valores de mesma dimensão, na matriz da área de busca, o melhor valor de casamento. A região que obtiver o melhor valor é considerada a possível localização do objeto alvo.

### 3.2 MODELOS DINÂMICOS LINEARES DISCRETOS

O rastreamento por Casamento de *Template* atua detectando o objeto em cada quadro de forma independente. Entretanto, é possível beneficiar-se de informações do movimento do objeto para melhorar o desempenho do rastreamento frente a alguns tipos de dificuldades.

Na tarefa de realizar o rastreamento visual de um objeto em vídeo, pode ser considerada a construção do modelo dinâmico do movimento do objeto. Consiste em ter, em cada quadro, um modelo para o movimento do objeto alvo além do conjunto de medidas de posição extraídas a partir de uma sequência de imagens. No rastreamento de objetos que utiliza o modelo dinâmico, as informações do objeto alvo são representadas em um vetor  $x_i$ , onde o índice  $i$  representa o  $i$ -ésimo quadro. O vetor  $x_i$  pode conter diversas informações do objeto alvo, por exemplo, a velocidade e a posição. A medida extraída por detecção no  $i$ -ésimo quadro é representada por  $y_i$ . A equação geral de atualização do estado atual do objeto para o próximo estado, ou seja, a atualização de  $x_i$  para  $x_{i+1}$  é escrita da seguinte forma:

$$x_{i+1} = f(x_i) + v_i,$$

$f(\cdot)$  é uma função de transição e  $v$  é o ruído do processo. O ruído do processo é necessário devido às incertezas existentes na natureza do movimento. Como exemplo, suponha que o modelo do movimento do objeto é da seguinte forma: a posição do objeto no passo  $i + 1$  é dada pela posição no passo  $i$  mais um ruído gaussiano. Nesse caso, a função  $f(\cdot)$  pode ser definida como

$$f(x_i) = d_i x_i,$$

sendo  $d_i$  uma matriz identidade que define o movimento do modelo. Se o estado do objeto é representado por suas coordenadas (a,b) na imagem, então o estado do objeto no quadro  $i + 1$  é obtido por

$$x_{i+1} = \begin{bmatrix} a_{i+1} \\ b_{i+1} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \times \begin{bmatrix} a_i \\ b_i \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} v_i.$$

O vetor de estados  $x_i$  é  $\begin{bmatrix} a_i \\ b_i \end{bmatrix}$ ,  $a_i$  e  $b_i$  representam as coordenadas do objeto no  $i$ -ésimo quadro.

A equação geral da medição pode ser definida assim:

$$y_i = g(x_i) + w_i,$$

$g(\cdot)$  é a função de medição e  $w$  é o ruído de medição. Em rastreamento de objetos em imagens uma medição é qualquer processo de detecção. Considerando o exemplo anterior, e que a função  $g(\cdot)$  pode ser definida como

$$g(x_i) = m_i x_i.$$

A medição é obtida assim:

$$y_i = \begin{bmatrix} a_i \\ b_i \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \times \begin{bmatrix} a_i \\ b_i \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} w_i.$$

Sendo  $m_i$ , neste exemplo, uma matriz identidade que define a observação. Este é o modelo mais simples possível, no entanto, dificilmente poderá ser útil, pelo fato de que nesta representação o objeto alvo aparenta estar num estado estacionário.

Um outro modelo do movimento é considerar que a velocidade do objeto alvo é constante. Neste modelo a posição  $(a_i, b_i)$  será a posição anterior mais o produto da velocidade pelo período de amostragem, acrescido pelo ruído gaussiano, ou seja,  $p_i = p_{i-1} + (\Delta t)v_{i-1} + w_i$ , onde o estado do objeto alvo agora é representado por  $p$  e  $v_i = v_{i-1}$  a velocidade constante. O modelo pode ser definido da seguinte forma:

$$\begin{bmatrix} p_{i+1} \\ v \end{bmatrix} = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix} \times \begin{bmatrix} p_i \\ v \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} w_i.$$

Além do modelo que considera a velocidade constante, pode ser utilizado também o modelo que considera a aceleração constante. Neste modelo a posição do objeto alvo é dada pela posição anterior mais o produto da velocidade pelo período de amostragem, acrescido pelo ruído gaussiano, onde a velocidade é igual a velocidade anterior mais a aceleração  $(a)$ . O modelo pode ser representado por:

$$\begin{bmatrix} p_{i+1} \\ v_{i+1} \\ a \end{bmatrix} = \begin{bmatrix} 1 & \Delta t & 0 \\ 0 & 1 & \Delta t \\ 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} p_i \\ v_i \\ a \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} w_i.$$

Pode ser adotado para a representação do modelo do objeto alvo modelos simples até modelos mais complexos, no entanto, modelos simples exigem menor capacidade computacional, porém são menos acurados. Uma observação importante na hora da criação do modelo é que quase sempre não se conhece o movimento do objeto alvo, pois se fosse conhecido não faria sentido algum à elaboração de um mecanismo de rastreamento visual de objetos em vídeo.

Um caso especial no rastreamento de objeto é quando a função  $f(\cdot)$ , da equação de atualização, e a função  $g(\cdot)$ , da equação da medição, são funções lineares e  $v_i$  e  $w_i$  são ruídos normais (gaussianos) e independentes. Neste caso, como a soma e o produto de variáveis aleatórias normais resulta em uma variável aleatória normal, todas as variáveis aleatórias geradas pelas equações de atualização e de medição serão normais (gaussianas). Será mostrado na seção 3.4 que esse caso tem solução fechada: o Filtro de Kalman.

### 3.3 RASTREAMENTO COMO UM PROBLEMA DE INFERÊNCIA

O problema de rastrear um objeto ao longo de uma sequência de quadros de um vídeo pode ser tratado como um problema de inferência probabilística, tendo como técnica chave a representação precisa da posição posterior do objeto alvo dado algumas medidas (detecção na imagem).

A tarefa de realizar rastreamento de objetos por inferência é realizado em duas etapas: uma etapa de predição e uma etapa de filtragem. Na primeira etapa as estimativas chamadas de preditoras ou *a priori*, são obtidas quando se tem um conjunto de medidas até  $i-1$  passo e deseja-se obter uma estimativa do vetor de estado no passo  $i$ . Na próxima etapa, a etapa de filtragem, com a observação no passo  $i$  e o valor encontrado no processo de predição, pode-se corrigir a estimativa do vetor de estado no passo  $i$ , processo também chamado de correção, e finalmente se obtém a medida corrigida no passo  $i$ , chamado valor *a posteriori*. Nessa tarefa são usados dois conjuntos de equações, as equações preditoras e as equações corretoras.

A modelagem de um objeto alvo pode ser representada da seguinte forma:

$X_i$ : é uma variável aleatória que representa o estado interno do objeto no  $i$ -ésimo quadro; para representar um determinado valor será usado  $x_i$  ( $x$  minúsculo).

$Y_i$ : é uma variável aleatória que representa as medidas no  $i$ -ésimo quadro; será usado  $y_i$  para denotar o valor de uma medição.

Algumas das questões que devem ser resolvidas em um rastreamento por inferência probabilística, são:

**Predição:** dado  $y_0, \dots, y_{i-1}$ , o objetivo é prever o estado do objeto no quadro  $i$ . A solução deste problema é a obtenção de uma representação de  $P(X_i|Y_0 = y_0, \dots, Y_{i-1} = y_{i-1})$ .

**Correção:** após ter uma medição por detecção  $y_i$ , é necessário computar uma representação de  $P(X_i|Y_0 = y_0, \dots, Y_i = y_i)$ . Esta é a distribuição do estado no quadro  $i$ .

## Suposições de independência

Para facilitar e simplificar o rastreamento devemos ter as seguintes suposições:

- Apenas o passado imediato importa, ou seja,  $P(X_i|X_1, \dots, X_{i-1}) = P(X_i|X_{i-1})$ .
- Medidas dependem apenas do estado atual, logo  $P(Y_i, Y_j, \dots, Y_k|X_i) = P(Y_i|X_i)$

$P(Y_j, \dots, Y_k|X_i)$ .

Essas duas suposições denotam que o problema de rastreamento de objetos por inferência tem o **Modelo de Markov Oculto**. A Figura 6 mostra graficamente a dependência entre os estados e as medições.

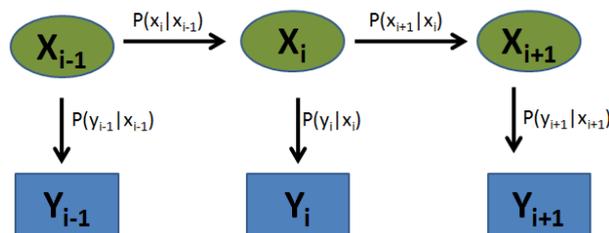


Figura 6 – Diagrama do rastreamento como um problema de inferência sequencial.

## Rastreamento por inferência

De forma resumida, o rastreamento por inferência consiste em calcular recursivamente um grau de confiança do estado  $x_i$  num tempo  $i$  dados os valores de  $y_0, y_1, \dots, y_i$  obtidos até o momento. Dessa forma, uma função densidade de probabilidade  $P(x_i|y_0, \dots, y_i)$  deve ser construída.

Procedendo indutivamente assume-se que temos  $P(X_0)$ , o qual é o valor *a priori*. Para corrigir esse valor após obter a medição  $Y_0$  (ou  $y_0$ ), aplica-se:

$$\begin{aligned}
P(X_0|Y_0 = y_0) &= \frac{P(y_0|X_0)P(X_0)}{P(y_0)} \\
&= \frac{P(y_0|X_0)P(X_0)}{\int P(y_0|X_0)P(X_0)dX_0} \\
&= \alpha P(y_0|X_0)P(X_0)
\end{aligned}$$

Nestas equações aplicam-se o teorema de Bayes. De forma geral, a representação da densidade  $P(X_i|y_0, \dots, y_i)$  pode ser obtida em  $i$  recursivamente seguindo as seguintes etapas:

### **Predição**

Seja a predição representada por:

$$P(x_i|y_0, \dots, y_{i-1}),$$

e assumamos que é conhecida a representação de  $P(x_{i-1}|y_0, \dots, y_{i-1})$ . Com as suposições de independência é possível escrever (FORSYTH; PONCE, 2002):

$$\begin{aligned}
P(X_i|y_0, \dots, y_{i-1}) &= \int P(X_i, X_{i-1}|y_0, \dots, y_{i-1})dX_{i-1} \\
&= \int P(X_i|X_{i-1}, y_0, \dots, y_{i-1})P(X_{i-1}|y_0, \dots, y_{i-1})dX_{i-1} \\
&= \int P(X_i|X_{i-1})P(X_{i-1}|y_0, \dots, y_{i-1})dX_{i-1}
\end{aligned}$$

### **Correção**

A correção busca calcular uma representação de

$$P(x_i|y_0, \dots, y_i),$$

que é obtida utilizando a resultante da etapa anterior. Novamente, com as suposições de independência é possível escrever (FORSYTH; PONCE, 2002):

$$\begin{aligned}
P(X_i|y_0, \dots, y_i) &= \frac{P(X_i, y_0, \dots, y_i)}{P(y_0, \dots, y_i)} \\
&= \frac{P(y_i|X_i, y_0, \dots, y_{i-1})P(X_i|y_0, \dots, y_{i-1})P(y_0, \dots, y_{i-1})}{P(y_0, \dots, y_i)} \\
&= P(y_i|X_i)P(X_i|y_0, \dots, y_{i-1}) \frac{P(y_0, \dots, y_{i-1})}{P(y_0, \dots, y_i)} \\
&= \frac{P(y_i|X_i)P(X_i|y_0, \dots, y_{i-1})}{\int P(y_i|X_i)P(X_i|y_0, \dots, y_{i-1})dX_i}
\end{aligned}$$

## 3.4 RASTREAMENTO UTILIZANDO FILTRO DE KALMAN

Um caso especial importante é quando, nos modelos que foram descritos, as densidades de probabilidade condicional são normais. Isso quer dizer que as variáveis envolvidas no modelo dinâmico são normais e que as funções descritas no modelo dinâmico são lineares, ou seja,  $P(x_i|y_0, \dots, y_{i-1})$  e  $P(x_i|y_0, \dots, y_i)$  são normais. Essa situação é um caso especial e a solução

ótima para esse caso é o Filtro de Kalman (FORSYTH; PONCE, 2002). Sabendo que a distribuição de uma normal é perfeitamente representado pela sua média e desvio padrão (ou matriz de covariância para casos multidimensionais), isso significa que basta manter a representação da média e do desvio padrão para as fases de predição e correção. O Filtro de Kalman é um conjunto de equações recursivas onde as representações da média e do desvio padrão para a predição e estimativa em cada fase são atualizadas.

### 3.4.1 O Filtro de Kalman para um vetor de estado em 1D

O modelo dinâmico agora é representado por:

$$\begin{aligned}x_i &\sim N(d_i x_{i-1}; \sigma_{d_i}^2), \\y_i &\sim N(m_i x_i; \sigma^2 m_i).\end{aligned}$$

É necessário manter a representação de  $P(x_i|y_0, \dots, y_{i-1})$  e  $P(x_i|y_0, \dots, y_i)$ . Em cada caso, é preciso representar apenas a média e o desvio padrão, pois as distribuições são normais. A representação da média e do desvio padrão, nas etapas de predição  $(\bar{x}_i^-, \sigma_i^-)$  e correção  $(\bar{x}_i^+, \sigma_i^+)$  são:

#### Predição

A predição é dada por:

$$P(X_i|y_0, \dots, y_{i-1}) = \int P(X_i|X_{i-1})P(X_{i-1}|y_0, \dots, y_{i-1})dX_{i-1}.$$

Pode ser expresso, utilizando de algumas propriedades de integração que,  $\bar{x}_i^- = d_i \bar{x}_{i-1}^+$  e que  $(\sigma_i^-)^2 = \sigma_{d_i}^2 + (d_i \sigma_{i-1}^+)^2$ .

#### Correção

A correção é dada por:

$$\begin{aligned}P(X_i|y_0, \dots, y_i) &= \frac{P(y_i|X_i)P(X_i|y_0, \dots, y_{i-1})}{\int P(y_i|X_i)P(X_i|y_0, \dots, y_{i-1})dX_i} \\ &= \alpha P(y_i|X_i)P(X_i|y_0, \dots, y_{i-1}).\end{aligned}$$

Sabendo que  $\bar{x}_{i-1}^-$  e  $\sigma_i^-$  representam  $P(x_i|y_0, \dots, y_{i-1})$  e por algumas propriedades de integração e com correspondência de padrão, então:

$$\begin{aligned}\bar{X}_i^+ &= \frac{\bar{X}_i^- \sigma_{m_i}^2 + m_i y_i (\sigma_i^-)^2}{\sigma_{m_i}^2 + m_i^2 (\sigma_i^-)^2}, \\ \sigma_i^+ &= \sqrt{\frac{\sigma_{m_i}^2 (\sigma_i^-)^2}{\sigma_{m_i}^2 + m_i^2 (\sigma_i^-)^2}}.\end{aligned}$$

### 3.4.2 O Filtro de Kalman para vetor de estado geral

Para a realização do rastreamento de um objeto em uma dimensão (1D) não foi necessário realizar nenhum cálculo com integrais. As mesmas técnicas abordadas para (1D) podem ser aplicadas para mais dimensões. A seguir são mostradas as atualizações das equações do Filtro de Kalman para qualquer dimensão:

#### Modelo dinâmico:

$$\begin{aligned}x_i &\sim N(D_i x_{i-1}, \Sigma_{d_i}), \\y_i &\sim N(M_i x_i, \Sigma_{m_i}).\end{aligned}$$

#### Suposições iniciais:

$\bar{x}_0^-$  e  $\Sigma_0^-$  são conhecidos

#### Predição:

$$\begin{aligned}\bar{x}_i^- &= D_i \bar{x}_{i-1}^+ \\ \Sigma_i^- &= \Sigma_{d_i} + D_i \Sigma_{i-1}^+ D_i\end{aligned}$$

#### Correção:

$$\begin{aligned}k_i &= \Sigma_i^- M_i^T [M_i \Sigma_i^- M_i^T + \Sigma_{m_i}]^{-1} \\ \bar{x}_i^+ &= \bar{x}_i^- + k_i [y_i - M_i \bar{x}_i^-] \\ \Sigma_i^+ &= [Id - k_i M_i] \Sigma_i^-\end{aligned}$$

O processo realizado pelo Filtro de Kalman pode ser mostrado conforme a Figura 7:



<b>Atualização de tempo("Predição")</b>	<b>Atualização da medida("Correção")</b>
Projeção do estado <i>a priori</i> $\bar{x}_i^- = D_i \bar{x}_{i-1}^+$ Projeção da covariância do erro <i>a priori</i> $\Sigma_i^- = \Sigma_{d_i} + D_i \Sigma_{i-1}^+ D_i$	Calcula o ganho de Kalman $k_i = \Sigma_i^- M_i^T [M_i \Sigma_i^- M_i^T + \Sigma_{m_i}]^{-1}$ Atualiza a estimativa $\bar{x}_i^+ = \bar{x}_i^- + k_i [y_i - M_i \bar{x}_i^-]$ Atualização da covariância do erro $\Sigma_i^+ = [Id - k_i M_i] \Sigma_i^-$



Figura 7 – Etapas do Filtro de Kalman (adaptado de (BISHOP; WELCH, 2001)).

Estimativa inicial:  $\bar{x}_0^-$  e  $\Sigma_0^-$

Após definidos os valores iniciais é realizada a predição, calculada a projeção do estado *a priori* e a projeção da covariância do erro *a priori*, representado pelo quadro **Atualização de tempo (“Predição”)**. Depois de realizada a predição é calculado o ganho de Kalman, o vetor de estados e a matriz de covariância, representado pelo quadro **Atualização da medida (“Correção”)**. O processo se repete para uma nova predição do vetor de estados. Por este motivo é dito que o Filtro de Kalman é um algoritmo recursivo, pois em cada período de tempo ao longo do processo como um todo, ele calcula as estimativas com base nas estimativas obtidas no período anterior (BISHOP; WELCH, 2001).

Observe que as equações de correção envolvem a inversão da matriz de covariância  $\Sigma_i$ . Na prática essa inversão pode apresentar problemas numéricos. Esse problema não será tratado aqui uma vez que o foco deste trabalho é aplicação e para implementação foram utilizados resolvidores de biblioteca, no caso MATLAB.

### 3.5 RASTREAMENTO UTILIZANDO FILTRO DE PARTÍCULAS

Para casos gerais do modelo dinâmico, onde as funções do modelo dinâmico são não lineares e as variáveis envolvidas não são normais, a principal dificuldade no rastreamento

é na manutenção de uma representação satisfatória da  $P(x_i|y_0, \dots, y_i)$ . Não existe uma solução ótima para esse problema, no entanto, existem abordagens que são úteis em muitas aplicações. Uma delas é o Filtro de Partículas, que foi desenvolvido originalmente para rastrear objetos em ambientes ruidosos (NUMMIARO et al., 2003).

Serão apresentados dois algoritmos que implementam o Filtro de Partículas, um algoritmo básico chamado de amostragem por importância sequencial ( Sequential Importance Sampling - SIS) e outro chamado de re-amostragem por importância sequencial (Sequential Importance Resampling - SIR) (RISTIC et al., 2004).

### 3.5.1 Sequential Importance Sampling (SIS)

O algoritmo *Sequential Importance Sampling* (SIS) utiliza o princípio da amostragem por importância (*Importance Sampling*). O princípio de amostragem por importância baseia-se em que dada uma distribuição  $p(x) \propto \pi(x)$  e  $x^k \approx z(x), k = 1, \dots, N$  amostras extraídas de uma distribuição  $z(\cdot)$ , chamada *densidade de importância*, é possível obter uma aproximação ponderada da densidade  $p(\cdot)$  a partir de:

$$p(x) \approx \sum_{k=1}^N w^k \delta(x - x^k),$$

$$w^k \propto \frac{\pi(x^k)}{z(x^k)},$$

$\delta(\cdot)$  é a função delta de *Dirac* (PFEIFFER, 2004). Na Figura 8 é ilustrado o princípio da amostragem por importância, sendo a função de densidade de probabilidade alvo (azul) aproximada por um conjunto de partículas (preto) extraídas utilizando uma densidade auxiliar (vermelho) conhecida, aqui definida como Gaussiana.

O princípio fundamental do algoritmo SIS é representar a função de densidade de probabilidade *a posteriori* utilizando partículas. Seja  $\{x_{0:i}^k, w_i^k\}_{k=1}^{N_s}$  a caracterização de uma função de densidade de probabilidade  $p(x_{0:i}|y_{1:i})$  onde  $\{x_{0:i}^k, k = 1, \dots, N_s\}$  é um conjunto de pontos associados a pesos  $\{w_i^k, k = 1, \dots, N_s\}$  e  $x_{0:i} = \{x_j, j = 0, \dots, i\}$  o conjunto de todos os estados até o instante  $i$ . Os pesos são normalizados de tal forma que  $\sum_{k=1}^{N_s} w_i^k = 1$ . Se as partículas  $x_{0:i}^k$  forem extraídas de uma *densidade de importância*  $z(x_{0:i}|y_{1:i})$ , a função de densidade de probabilidade *a posteriori* em  $i$  pode ser aproximada como:

$$p(x_{0:i}|y_{1:i}) \approx \sum_{k=1}^{N_s} w_i^k \delta(x_{0:i} - x_{0:i}^k),$$

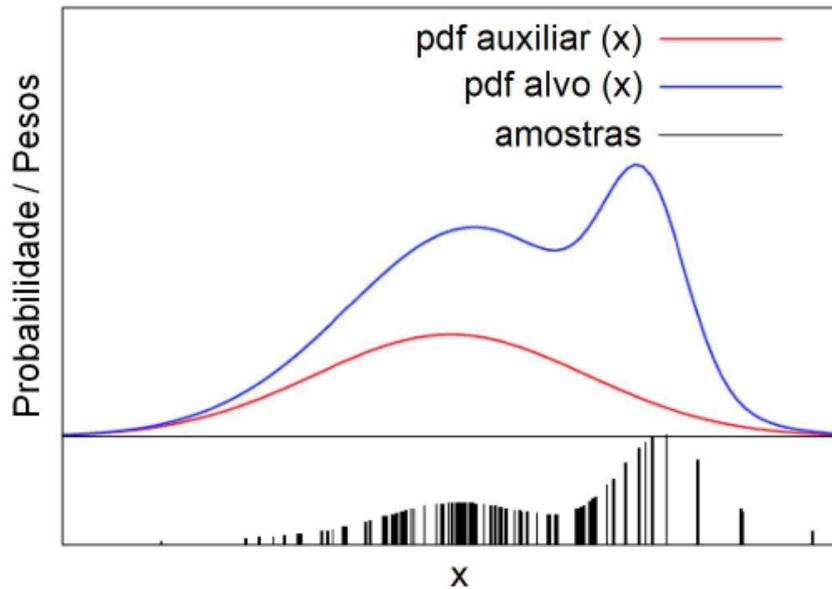


Figura 8 – Princípio da amostragem por importância (Adaptado de (GRISSETTI; STACHNISS, 2007)).

$$w_i^k \propto \frac{p(x_{0:i}^k | y_{1:i})}{z(x_{0:i}^k | y_{1:i})},$$

A etapa da **evolução do modelo** consiste em evoluir cada uma das partículas existentes  $x_{0:i-1}^k \approx z(x_{0:i-1} | y_{1:i-1})$  para o novo estado  $x_i^k \approx z(x_i | x_{0:i-1}, y_{1:i})$ . Se puder escrever a densidade de importância escolhida como:

$$z(x_{0:i} | y_{1:i}) = z(x_i | x_{0:i-1}, y_{1:i}) z(x_{0:i-1} | y_{1:i-1}),$$

então será possível obter as novas amostras  $x_{0:i}^k \approx z(x_{0:i} | y_{1:i})$ .

O passo seguinte é o de **atualizar** a predição do sistema. Nesse passo são determinados os pesos das amostras  $w_i^k$  no instante  $i$ . Desenvolvendo  $p(x_{0:i} | y_{1:i})$ , obtém-se:

$$p(x_{0:i} | y_{1:i}) = p(y_i | x_i) p(x_i | x_{i-1}) p(x_{0:i-1} | y_{1:i-1}).$$

Considerando que a densidade de importância é dependente apenas de  $x_{i-1}$  e  $y_i$ , ou seja,  $z(x_i | x_{0:i-1}, y_{1:i}) = z(x_i | x_{0:i-1}, y_i)$ , a função de densidade de probabilidade atualizada é aproximada por:

$$p(x_i | y_{1:i}) \approx \sum_{k=1}^{N_s} w_i^k \delta(x_i - x_i^k),$$

$$w_i^k \propto w_{i-1}^k \frac{p(y_i | x_i^k) p(x_i^k | x_{i-1}^k)}{z(x_i^k | x_{i-1}^k, y_i)}.$$

### 3.5.2 Sequential Importance Resampling (SIR)

No algoritmo SIS pode acontecer que depois de algumas iterações a maior parte das partículas possuam importância insignificante, levando a um grande esforço computacional na atualização de partículas não relevantes para a estimação da densidade *a posteriori*. A solução adotada para evitar este problema foi criar uma etapa adicional. Trata-se da etapa de **seleção** ou **re-amostragem**, gerando o algoritmo **Sequential Importance Resampling (SIR)**.

A técnica de Re-amostragem (*Resampling*) evita a degeneração das partículas. Esta técnica tenta eliminar as partículas de menor peso para que o filtro concentre os esforços nas “melhores” partículas, de tal forma que um novo conjunto  $\{x_i^{k*}, Ns^{-1}\}$  é obtido re-amostrando  $\{x_i^k, w_i^k\}$  com reposição. A re-amostragem pode ser aplicada em todas as iterações  $i$  do algoritmo ou quando o nível de degeneração atingir um limiar estabelecido.

### 3.5.3 Funcionamento qualitativo do algoritmo SIR

O algoritmo SIR é, portanto, o mesmo algoritmo SIS acrescido da etapa de re-amostragem.

O funcionamento qualitativo do algoritmo SIR é ilustrado na Figura 9, sendo dividido em três etapas: predição, correção e re-amostragem. As partículas são representadas no formato de círculos, e o tamanho do círculo representa o peso da partícula.

Seguindo na figura uma linha do tempo de cima para baixo, temos os seguintes passos:

Passo 1 (predição): as partículas são distribuídas, com pesos iguais, no espaço amostral.

Passo 2 (correção): com base no valor da observação, calculada, por exemplo, por alguma técnica de detecção, e na posição anterior da partícula, é estimada qual será a próxima posição para cada partícula e é calculado um novo peso de cada partícula. O peso é uma aproximação para a probabilidade do objeto alvo estar na posição prevista pela observação.

Passo 3 (re-amostragem): as partículas são redefinidas, no intuito de descartar partículas com baixo peso e explorar a região que as partículas possuem pesos elevados, de tal forma que vão surgir a mesma quantidade de partículas que existia no primeiro passo, porém em regiões mais próximas da posição prevista pela observação.

Fluxogramas detalhados dos algoritmos desta seção serão apresentados no capítulo

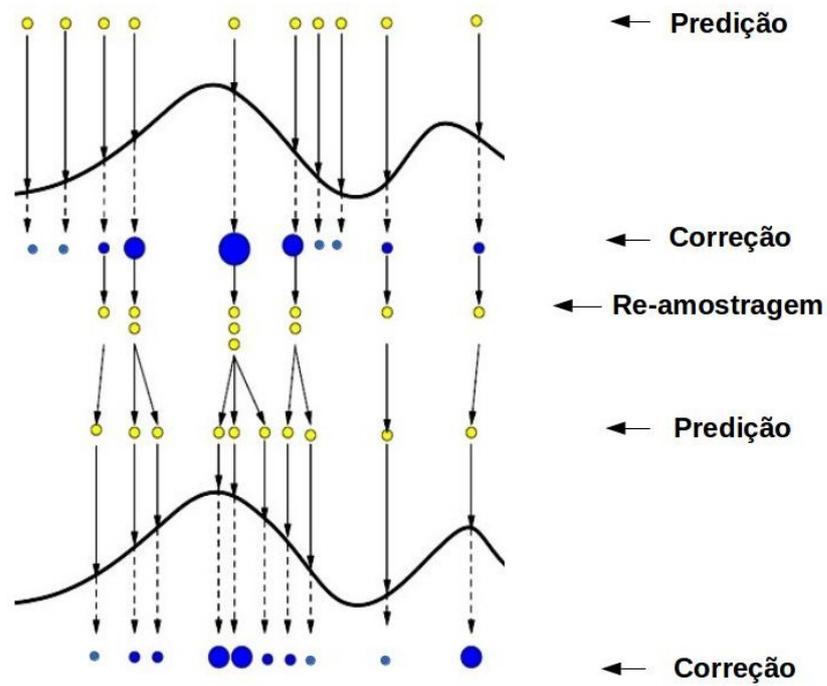


Figura 9 – Ilustração do funcionamento do Filtro de Partículas.

4.

### 3.6 CONCLUSÕES DO CAPÍTULO

Neste capítulo foram apresentados os fundamentos teóricos dos três métodos de rastreamento comparados nesta dissertação. O método Casamento de *Template* que detecta o objeto em cada quadro de forma independente será utilizado como *baseline*. Estimação utilizando Filtro de Kalman verifica a validade de uma aproximação do ruído como gaussiano. O Filtro de Partículas é representativo de uma técnica que não impõe restrições sobre a natureza das distribuições envolvidas.

## 4 IMPLEMENTAÇÃO

A implementação dos rastreadores foi realizada em MATLAB. Além de algumas funções básicas de apoio do MATLAB, foram utilizadas as seguintes funções do *toolbox vision*:

- vision.TemplateMatcher.
- vision.KalmanFilter.

O Casamento de *Template* utilizando SAD, SSD ou Corr2 é muito simples de implementar. Entretanto, resolveu-se utilizar a função MATLAB para tornar o código mais compacto. Já a função vision.KalmanFilter foi utilizada para não ter que se preocupar com a estabilidade numérica da implementação.

Os algoritmos trabalham sobre imagens em nível de cinza e permitem a seleção de qualquer alvo como objeto de rastreamento. Para simplicidade de comparação foi considerado como objeto alvo em todos os vídeos selecionados o rosto em perfil de uma pessoa, e todas as implementações dos algoritmos permitem a seleção de apenas um objeto (ou seja, nas implementações realizadas não é possível rastrear múltiplos objetos).

Alguns passos de um rastreador para outro rastreador são iguais, sendo descrito apenas no rastreador com Casamento de *Template*.

### 4.1 SISTEMA RASTREADOR COM CASAMENTO DE *TEMPLATE*

Os passos da implementação do rastreamento feito com Casamento de *Template* ocorre conforme a Figura 10.

A primeira etapa é a seleção do objeto no primeiro quadro em que aparece o objeto alvo. O algoritmo implementado gera uma imagem inicial que aguarda a interação com usuário<sup>1</sup>, na qual ele deve selecionar o objeto alvo a ser rastreado. A seleção do objeto ocorre na forma de um retângulo para representar um objeto, não havendo reajuste de tamanho do retângulo durante a execução. Tudo que fica dentro do retângulo passa a corresponder ao modelo alvo do rastreamento. Na Figura 11 é mostrado um exemplo de seleção do objeto alvo.

Feita a seleção do objeto alvo no quadro  $i$ , no próximo quadro, quadro  $i + 1$ , é determinado uma área de busca em torno da área que o objeto alvo estava no quadro  $i$ . A área de

<sup>1</sup> A interação ocorre da seguinte forma: é mostrado uma tela com o primeiro quadro que aparece o objeto alvo, em seguida, com o mouse, o usuário seleciona o objeto alvo desenhando um retângulo.

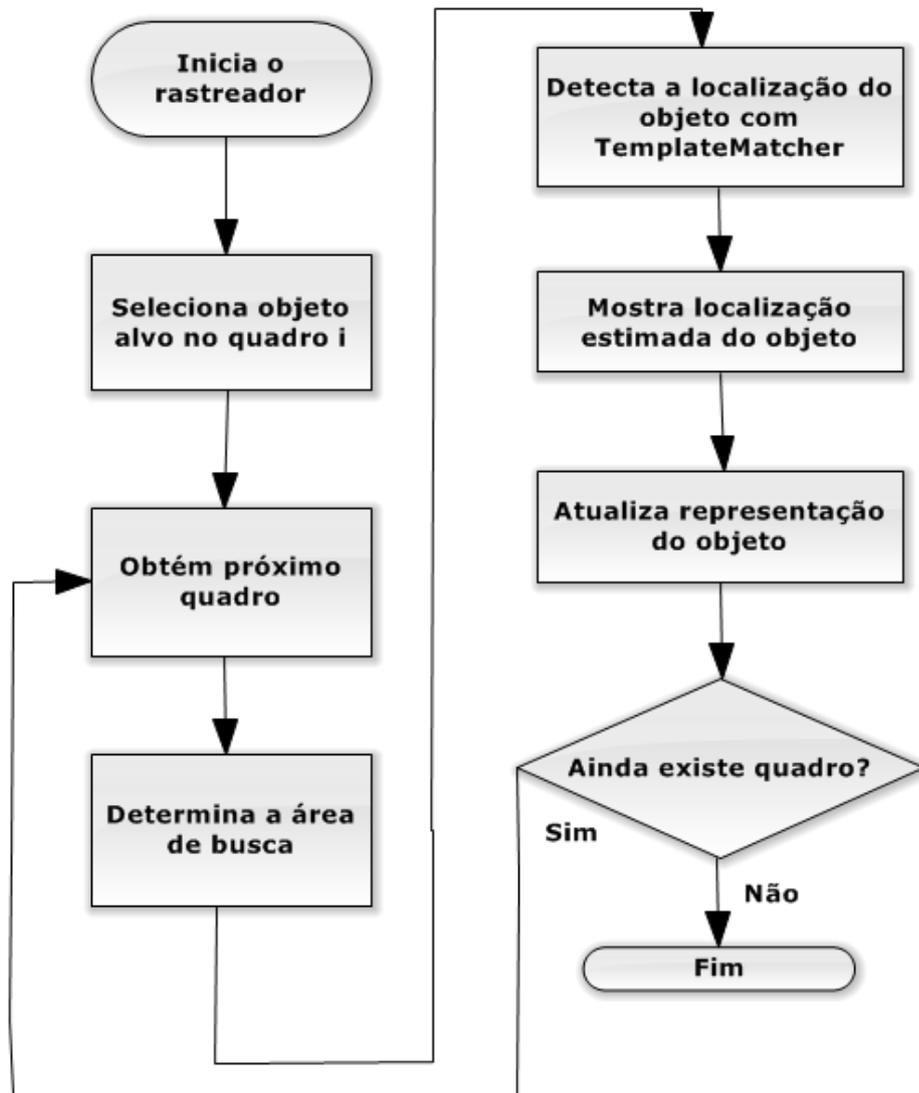


Figura 10 – Fluxograma de rastreamento com Casamento de *Template*.



Figura 11 – Seleção do rosto em perfil de uma pessoa como objeto alvo.

busca pode ser determinada como sendo todo o quadro ou como sendo parte dele. O tamanho da área de busca é um fator determinante no tempo de execução de um rastreador, na Tabela 1 é mostrado o tempo de execução considerando a área de busca o quadro todo (testado com um quadro do vídeo 1, na qual possui tamanho 320x240), 10 e 30 *pixels* a mais de cada lado da área do objeto alvo .

Tabela 1 – Tempo de execução definidos pelo tamanho da área de busca.

	10 pixels + área do objeto alvo	30 pixels + área do objeto alvo	Quadro todo
Tempo (s)	0,0672	0,3349	4,5703

Levando em consideração que a natureza do movimento do objeto rastreado é suave e o tempo de execução, foi definido que a área de busca possui dimensão da área do objeto alvo mais 10 *pixels* de cada lado.

O próximo passo é a detecção da localização do objeto alvo no quadro  $i + 1$ . O rastreador usa a classe **vision.TemplateMatcher** da biblioteca do MATLAB. Para realizar a localização do objeto alvo os rastreadores foram testados com dois critérios de casamento, a correlação entre matrizes (Corr2) e a soma das diferenças absoluta (SAD), no entanto, optou-se por utilizar como critério de casamento a soma das diferenças absolutas pois o número de operações a serem executadas são menores do que os outros critérios apresentados, reduzindo assim o esforço computacional. A classe **vision.TemplateMatcher** devolve como valor um ponto central da região da área de busca que possui a melhor similaridade entre as regiões da área de busca com a região que representa o objeto alvo.

Depois de encontrada as coordenadas da possível localização do objeto alvo (o ponto central da região), no passo **Mostra localização estimada do objeto**, é inserido um retângulo na posição indicada de mesma dimensão do objeto alvo, para mostrar o resultado do rastreamento.

Antes do próximo quadro é atualizada a representação do objeto alvo, sendo definida de forma empírica, como sendo a representação do objeto alvo em cada quadro constituída em 20% da representação do objeto alvo selecionado na primeira imagem e mais 80% da representação do objeto rastreado na imagem atual.

## 4.2 SISTEMA RASTREADOR COM FILTRO DE KALMAN

A implementação do rastreador feito com Filtro de Kalman ocorre conforme a Figura 12.

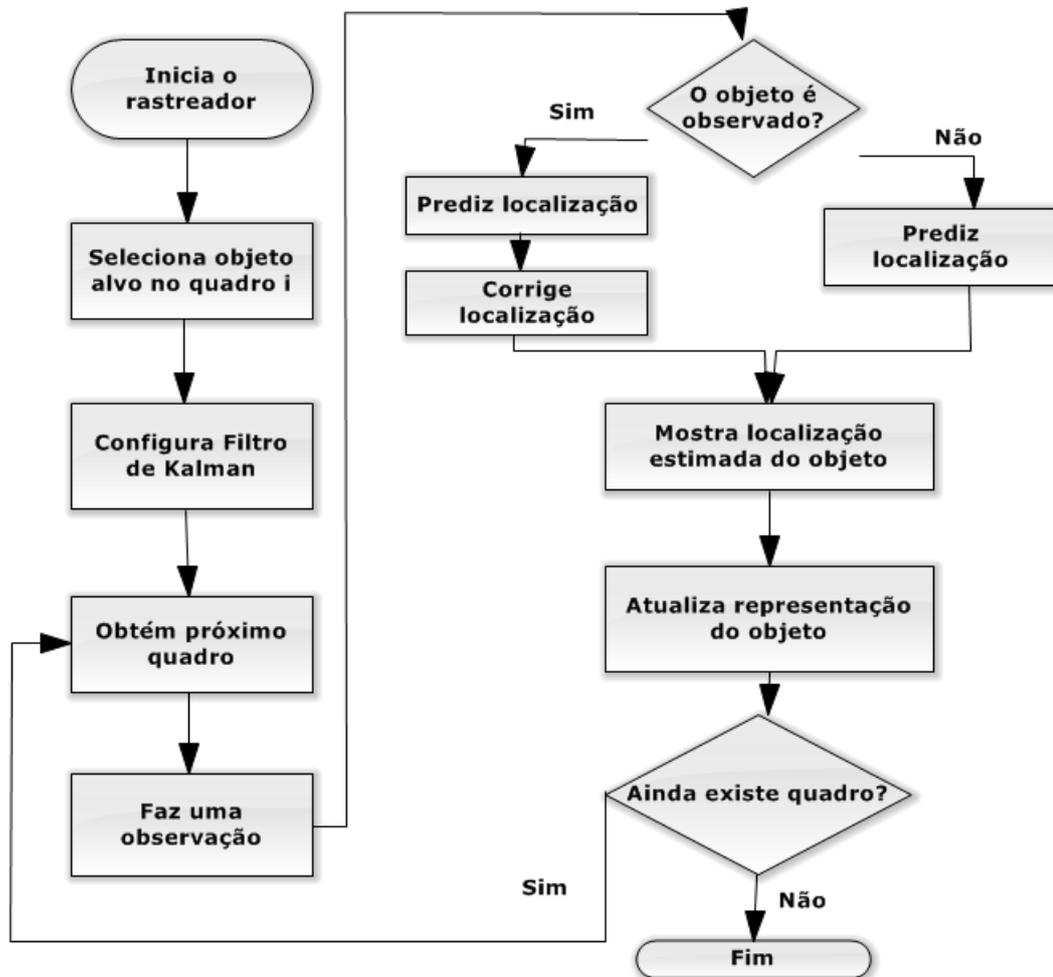


Figura 12 – Fluxograma de rastreamento com Filtro de Kalman.

Os passos **Seleciona objeto alvo no quadro i**, **Mostra localização estimada do objeto** e **Atualiza representação do objeto** são iguais aos descritos no rastreador com Casamento de *Template*.

No passo **Configura Filtro de Kalman** utiliza a classe **vision.KalmanFilter**. Para ela é passado como parâmetro o modelo dinâmico do objeto. O modelo é configurado para considerar que a velocidade do objeto alvo é constante. No modelo, o estado é representado por:

$$\begin{bmatrix} a_i \\ b_i \\ v_x \\ v_y \end{bmatrix}$$

onde  $a_i$  e  $b_i$  representam as coordenadas do objeto alvo no instante  $i$ , tendo como valor inicial extraído da seleção do objeto alvo realizado pelo usuário,  $v_x$  representa a velocidade na coordenada  $x$  e  $v_y$  a velocidade na coordenada  $y$ , ambos constantes e iguais a 1. A matriz de transição do modelo do estado utilizado foi:

$$\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

e a matriz do modelo da medição é dada por:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

Foram definidas de forma empírica que a variância do ruído do processo é 0,1 e do ruído da medição é igual a 4.

Após a obtenção do próximo quadro é feita uma nova observação. A observação utiliza Casamento de *Template* utilizando o índice SAD e retorna a posição de distância mínima. Nesse rastreamento, o objeto só é considerado detectado se SAD estiver abaixo de um certo limiar. Se uma SAD tão baixo não é obtido o objeto é considerado não detectado e o rastreamento baseia-se apenas no passo de predição.

No passo de verificação se o objeto foi observado, sendo considerado que o objeto é observado se o valor da soma das diferenças absoluta for menor do que 40000 (definido de forma empírica), caso contrário é considerado que o objeto não é observado, é realizado a predição do estado do objeto alvo e caso o objeto seja considerado observado, além da predição, é feita a correção da predição do estado do objeto alvo. Para realizar a predição e a correção foram chamado os métodos **predict** e **correct**, da classe **vision.KalmanFilter**.

Após a etapa de verificação se o objeto alvo foi observado ou não, a classe **vision.KalmanFilter** retorna as coordenadas da possível localização do objeto alvo. Com as

coordenadas da possível localização do objeto é mostrada a localização estimada do objeto através de um retângulo em torno delas.

O passo de atualização da representação do objeto é realizado e em seguida é verificado se ainda existe quadros. Se sim, o algoritmo volta para o passo que obtém o próximo quadro; se não, termina.

#### 4.3 SISTEMA RASTREADOR COM FILTRO DE PARTÍCULAS

A implementação do rastreador com Filtro de Partículas ocorre conforme a Figura 13.

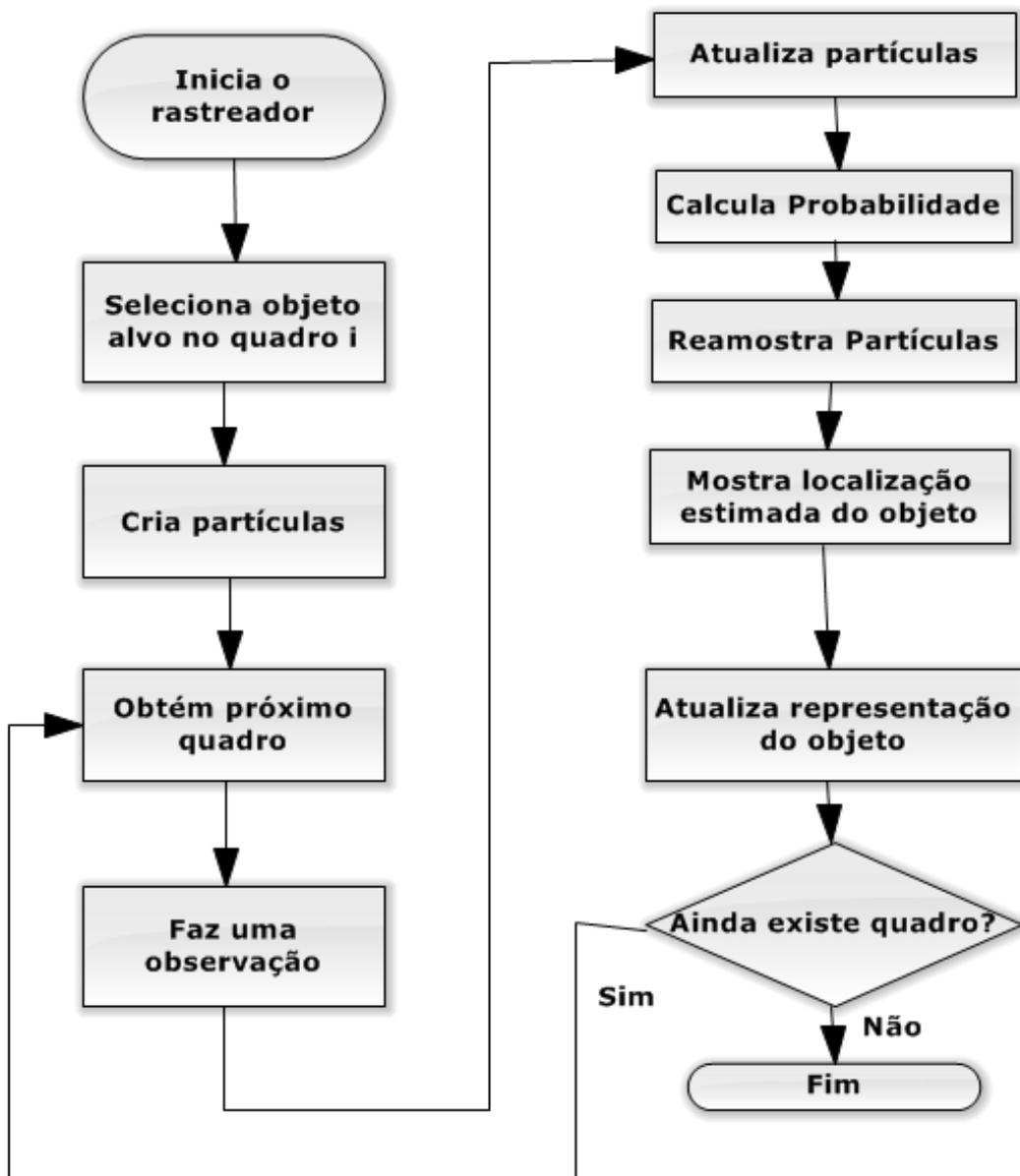


Figura 13 – Fluxograma de rastreamento com Filtro de Partículas.

Os passos **Seleciona objeto alvo no quadro i** e **Atualiza representação do objeto** são iguais aos descritos no rastreador com Casamento de *Template*.

No passo **Cria partículas** são criadas 4000 mil partículas, cada partícula é representada pela posição a, posição b, deslocamento X e deslocamento Y. A representação das 4000 partículas é feita através de uma matriz de 4 linhas e 4000 colunas, onde a primeira e segunda linhas representam a posição inicial das partículas no eixo x e y em *pixels*, respectivamente, com valores iniciais aleatórios<sup>2</sup> e a terceira e quarta linhas representam o deslocamento da partícula em *pixels* no eixo x e no eixo y e ambos são inicializados com valor zero.

No passo **Faz uma observação** foi utilizado o mesmo método do rastreador com Filtro de Kalman.

O próximo passo é a atualização das partículas. Neste passo as partículas são atualizadas de acordo com o modelo dinâmico do objeto, onde, no modelo, cada partícula é representado por:

$$\begin{bmatrix} a_i \\ b_i \\ \text{Deslocamento}X \\ \text{Deslocamento}Y \end{bmatrix}$$

O modelo do estado é representado por:

$$\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

E o modelo de medição é representado por:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

O ruído do processo e o ruído da medição são valores aleatórios gerados pela função *randn* do MATLAB. Essa função gera números aleatórios com distribuição normal de média zero e desvio padrão igual a um e é multiplicado por 25 para o ruído do processo e por 5 para o ruído da medição.

<sup>2</sup> Para o eixo x varia de 0 até o tamanho máximo do quadro no eixo x e para o eixo y varia de 0 até o tamanho máximo do quadro no eixo y.

A atualização das partículas consiste na soma da posição atual da partícula com o seu deslocamento acrescentando os ruídos.

No passo seguinte, **Calcula Probabilidade**, são calculados os pesos das partículas considerando o resultado da observação. Para o cálculo desta probabilidade foi considerado que as partículas com a posição mais próxima da observação possuem probabilidades maiores e as mais distantes probabilidades menores, com probabilidades inversamente proporcionais a distância.

Após as probabilidades das partículas já devidamente calculadas, é executado o passo **Reamostra Partículas**. A finalidade desse passo é deixar apenas as partículas com maior probabilidade (peso) sobreviverem. Nesse passo são definidas faixas de probabilidades que expressam a probabilidade que cada partícula tem de existir baseada na matriz calculada no passo anterior (a matriz que contém a probabilidade de cada partícula), e a partir das faixas definir uma distribuição aleatória. As faixas são definidas através da exponencial da subtração da probabilidade das partículas pelo maior valor de probabilidade da partícula, em seguida divide cada termo pelo somatório da mesma e faz a soma cumulativa. Partículas com probabilidades maiores possuem faixas maiores e por isso a distribuição aleatória concentra mais elementos nelas. Com o uso do método *histc* do MATLAB, é possível ter informação das faixas de probabilidade em que a distribuição se encontra, sendo reproduzidas somente as partículas representadas por essas faixas, ou seja, as partículas mais aptas a sobreviverem (com maior probabilidade) são reproduzidas novas partículas iguais a elas.

No passo **Mostra localização estimada do objeto** é apresentado as partículas no quadro, onde a localização das partículas expressam a localização estimada do objeto.

O passo de atualização da representação do objeto é realizado e em seguida é verificado se ainda existe quadros. Se sim, o algoritmo volta para o passo que obtém o próximo quadro; se não, termina.

#### 4.4 CONCLUSÕES DO CAPÍTULO

Este capítulo descreveu as implementações utilizando fluxogramas. A programação utilizou algumas funções essenciais previamente implementadas no MATLAB. Pode-se concluir que o trabalho de implementação tornou-se relativamente simples ao contar com o apoio destas funções.

## 5 METODOLOGIA DE AVALIAÇÃO

A metodologia de avaliação consiste em caracterizar os vídeos utilizados e definir as métricas de desempenho. Os vídeos utilizados para avaliar os rastreadores são listados na Tabela<sup>1</sup> 2. Eles estão disponíveis na internet e são amplamente utilizados em trabalhos de rastreamento visual. A seção 5.1 apresenta uma caracterização detalhada desses vídeos e a seção 5.2 define as métricas de desempenho utilizadas que são uma medida de acurácia e uma medida de robustez do rastreamento.

Tabela 2 – Vídeos utilizados para avaliar os rastreadores.

	Nº quadros	Taxa de quadros	Dimensões	Link para download do vídeo (acessados em 12/09/2014)
Vídeo 1	534	25 quadros/s	320x240	<a href="http://www.openvisor.org/video_Download.asp?idVideo=45">http://www.openvisor.org/video_Download.asp?idVideo=45</a>
Vídeo 2	839	25 quadros/s	368x272	<a href="http://www.openvisor.org/video_Download.asp?idVideo=120">http://www.openvisor.org/video_Download.asp?idVideo=120</a>
Vídeo 3	542	25 quadros/s	352x288	<a href="http://www.openvisor.org/video_Download.asp?idVideo=204">http://www.openvisor.org/video_Download.asp?idVideo=204</a>
Vídeo 4	172	25 quadros/s	352x288	<a href="http://box.vicos.se/vot/vot2013.zip">http://box.vicos.se/vot/vot2013.zip</a>

### 5.1 CARACTERIZAÇÃO DOS VÍDEOS

Os vídeos que foram selecionados ressaltam fatores importantes no rastreamento de um objeto em vídeo, como a Mudança de Aparência (MA), Oclusão (O), Manobra (M) e Mudança de Iluminação (MI) do objeto na sequência de quadros. Foram definidos e expressos na tabela abaixo os seguintes valores para MA, O, M e MI:

Tabela 3 – Valores para as características MA, O, M e MI.

	Nenhuma	Parcial ou Total
MA	0	1
O	0	1
M	0	1
MI	0	1

O rótulo "0" representa que não ocorreu nenhuma mudança de aparência ou nenhuma oclusão ou nenhuma manobra ou nenhuma mudança de iluminação; "1" representa que ocorreu

<sup>1</sup> O link para o vídeo 4 contém um conjunto de vídeos da competição VOT 2013 (Visual Object Tracking 2013), onde o vídeo utilizado neste trabalho foi o SUNSHADE.

uma parcial ou total mudança de aparência ou parcial ou total oclusão ou parcial ou total manobra ou parcial ou total mudança de iluminação.

Para cada vídeo é feito a análise de todos os seus quadros, mostrando se determinado quadro teve mudança de iluminação, manobra, mudança de aparência e oclusão do objeto alvo. Ainda na análise do vídeo é apresentada a trajetória real do objeto selecionado na sequência de quadros.

### 5.1.1 Análise do Vídeo 1

A sequência de quadros do vídeo 1 ilustra um homem que caminha em direção a uma cadeira, se senta, observa um cachorro passando em sua frente e em seguida se levanta e volta para onde veio. Nessa sequência de quadros o objeto alvo (a cabeça do homem) sofre variações de aparência e de manobra, no entanto, nenhuma oclusão e nenhuma mudança de iluminação. Para o vídeo 1 foram selecionados 350 quadros, onde o primeiro quadro <sup>2</sup> a ser analisado foi 111 e o último quadro o 460, os resultados da rotulação de acordo com a Mudança de Aparência (MA), Oclusão (O), Manobra (M) e Mudança de Iluminação (MI) são expressos na tabela abaixo.

Tabela 4 – Rotulação do vídeo 1 – MA, O, M e MI.

MA	0	1	1	1	1	1
O	0	0	0	0	0	0
M	0	1	0	1	0	1
MI	0	0	0	0	0	0
Trecho do vídeo	111/136	137/150	151/178	179/190	191/244	245/290

MA	1	1	1	1	1
O	0	0	0	0	0
M	0	1	0	1	0
MI	0	0	0	0	0
Trecho do vídeo	291/339	340/350	351/404	405/440	441/460

O gráfico <sup>3</sup> com a trajetória real do centro da região que representa o objeto alvo no vídeo é mostrado na Figura 14.

<sup>2</sup> Os vídeos 1,2 e 3 foram analisados a partir do quadro 111, pois o objeto alvo começou a aparecer a partir do quadro 110. O quadro 110 serviu para o usuário selecionar o objeto alvo.

<sup>3</sup> As setas indicam o sentido do objeto alvo no vídeo, onde a seta em azul indica o sentido inicial.

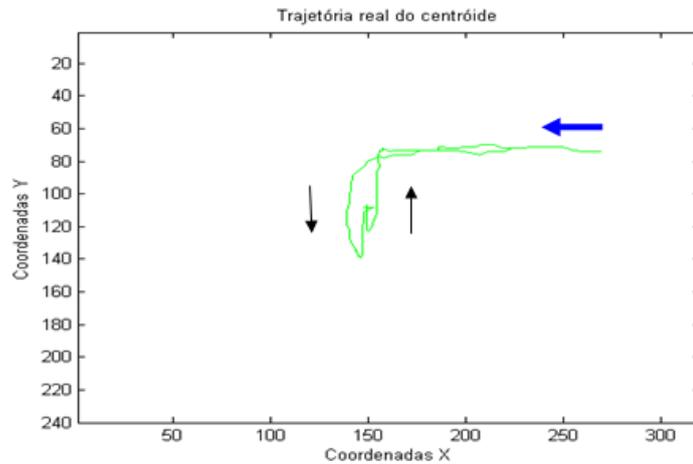


Figura 14 – Trajetória do centro do objeto alvo no vídeo 1.

Observe que a trajetória é simples, entretanto a trajetória do objeto alvo nas coordenadas do quadro muitas vezes não é o fator mais relevante. Em muitos vídeos o principal fator que torna o rastreamento difícil é a variação de *background*.

### 5.1.2 Análise do Vídeo 2

O vídeo 2 ilustra um homem entrando em um depósito, se movimentando em torno de algumas caixas, dando uma volta e depois se escondendo por trás delas, demorando alguns instantes e saindo. Nessa sequência o objeto alvo (a cabeça do homem) sofre variações de aparência, de manobra e oclusão. Nesse vídeo foram selecionados 650 quadros, onde o primeiro quadro a ser analisado foi o 111 e o último quadro o 760. Os resultados da rotulação de acordo com a Mudeça de Aparência (MA), Oclusão (O), Manobra (M) e Mudança de Iluminação (MI) são expressos na tabela abaixo.

Tabela 5 – Rotulação do vídeo 2 – MA, O, M e MI.

MA	0	0	1	1	1	1	1
O	0	1	1	0	0	0	0
M	0	0	0	1	0	1	0
MI	0	0	0	0	0	0	0
Trecho do vídeo	111/118	119/134	135/164	165/180	181/212	213/270	271/294

MA	1	1	1	0	0	1	1
O	0	0	1	1	0	0	0
M	1	0	0	0	0	1	0
MI	0	0	0	0	0	0	0
Trecho do vídeo	295/300	301/365	366/395	396/419	420/438	439/445	446/491

MA	1	1	1	1	1	0	0
O	1	0	0	1	0	0	0
M	0	1	0	1	0	1	0
MI	0	0	0	0	0	0	0
Trecho do vídeo	492/657	658/662	663/674	675/684	685/708	709/718	719/727

MA	1
O	0
M	0
MI	0
Trecho do vídeo	728/760

O gráfico <sup>4</sup> com a trajetória real do centro da região que representa o objeto alvo no vídeo é mostrado na Figura 15.

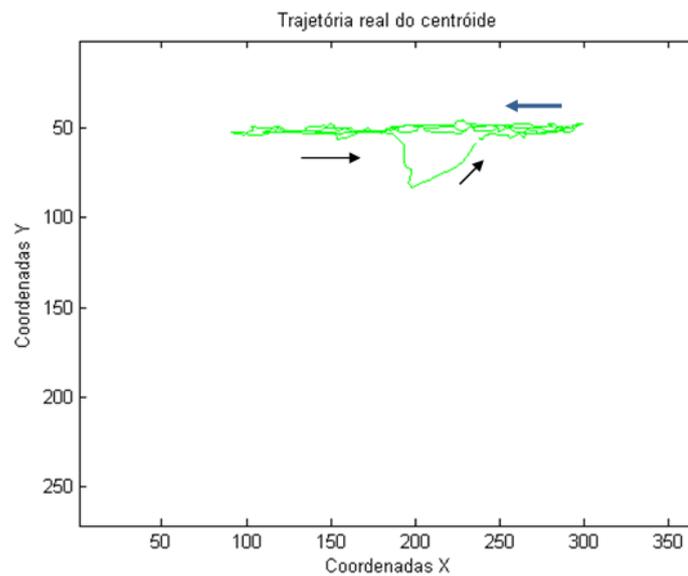


Figura 15 – Trajetória do centro do objeto alvo no vídeo 2.

<sup>4</sup> As setas indicam o sentido do objeto alvo no vídeo, onde a seta em azul indica o sentido inicial.

### 5.1.3 Análise do Vídeo 3

No vídeo 3 ilustra um homem entrando em uma sala, se movimentando de tal forma a se aproximar de umas lixeiras e depois indo para trás de uns quadros de texto, se movimentando por trás dos quadros e em seguida saindo da sala. Nessa sequência o homem, no início, realiza algumas manobras bruscas, e se movimenta de forma rápida pela sala, por esses motivos o objeto alvo (a cabeça do homem) sofre diversas variações de aparência, de manobra e passa por um longo período de oclusão. Nesse vídeo foram selecionados 410 quadros, o primeiro quadro começa no 111 e o último quadro é o 520, os resultados da rotulação dos quadros de acordo com a Mudeza de Aparência (MA), Oclusão (O), Manobra (M) e Mudança de Iluminação (MI) são expressos na tabela abaixo.

Tabela 6 – Rotulação do vídeo 3 – MA, O, M e MI.

MA	0	0	1	1	1	1	1
O	0	0	0	0	0	0	0
M	0	1	1	0	1	0	1
MI	0	0	0	0	0	0	0
Trecho do vídeo	111/126	127/130	131/137	138/156	157/189	190/191	192/205

MA	1	1	1	0	1	1	0
O	0	0	0	0	0	1	1
M	0	1	0	0	0	0	0
MI	0	0	0	0	0	0	0
Trecho do vídeo	206/241	242/260	261/265	266/288	289/299	300/421	422/462

MA	1	1	1	1	1	1	1
O	1	0	0	0	0	0	0
M	0	0	1	0	1	1	0
MI	0	0	0	0	0	0	0
Trecho do vídeo	463/466	467/469	470/476	477/494	495/505	506/510	511/520

O gráfico<sup>5</sup> com a trajetória real do centro da região que representa o objeto alvo no vídeo é mostrado na Figura 16.

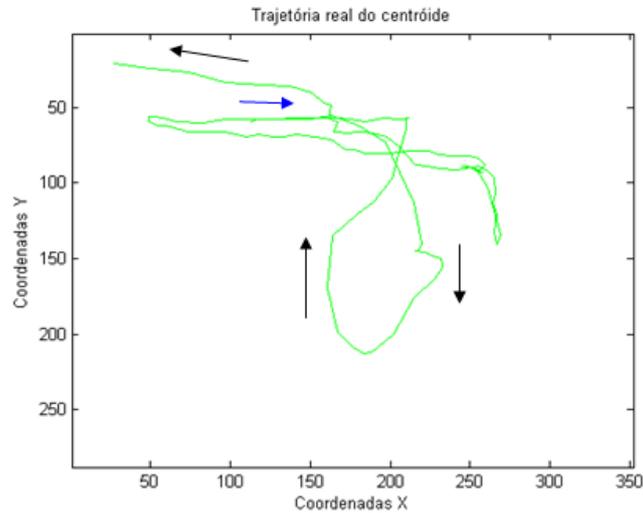


Figura 16 – Trajetória do centro do objeto alvo no vídeo 3.

#### 5.1.4 Análise do Vídeo 4

O vídeo 4 ilustra um homem se movimentando de um lado para outro entre regiões com diferentes níveis de iluminação. Nessa sequência o objeto alvo (a cabeça do homem) sofre por diversas vezes mudança de iluminação e se movimenta de forma rápida de um lado para outro, no entanto, não sofre nenhuma oclusão, nem manobra e nem mudança de aparência. Nesse vídeo foram selecionados 171 quadros, o primeiro quadro começa no 2 e o último quadro é o 172, os resultados da rotulação dos quadros de acordo com a Mudança de Aparência (MA), Oclusão (O), Manobra (M) e Mudança de Iluminação (MI) são expressos na tabela abaixo.

Tabela 7 – Rotulação do vídeo 4 – MA, O, M e MI.

MA	0	0	0	0	0	0	0
O	0	0	0	0	0	0	0
M	0	0	0	0	1	0	1
MI	0	1	0	1	1	1	1
Trecho do vídeo	2/6	7/18	19/33	34/48	49/51	52/57	58/59

<sup>5</sup> As setas indicam o sentido do objeto alvo no vídeo, onde a seta em azul indica o sentido inicial.

MA	0	0	0	0	0	0	0
O	0	0	0	0	0	0	0
M	0	0	0	1	0	0	0
MI	1	0	1	1	1	0	1
Trecho do vídeo	60/66	67/76	77/81	82/83	84/85	86/95	96/97

MA	0	0	0	0	0	0	0
O	0	0	0	0	0	0	0
M	1	0	0	1	0	0	0
MI	1	1	0	0	0	1	0
Trecho do vídeo	98/99	100/102	103/129	130/131	132/151	152/165	166/172

O gráfico <sup>6</sup> com a trajetória real do centro da região que representa o objeto alvo no vídeo é mostrado na Figura 17.

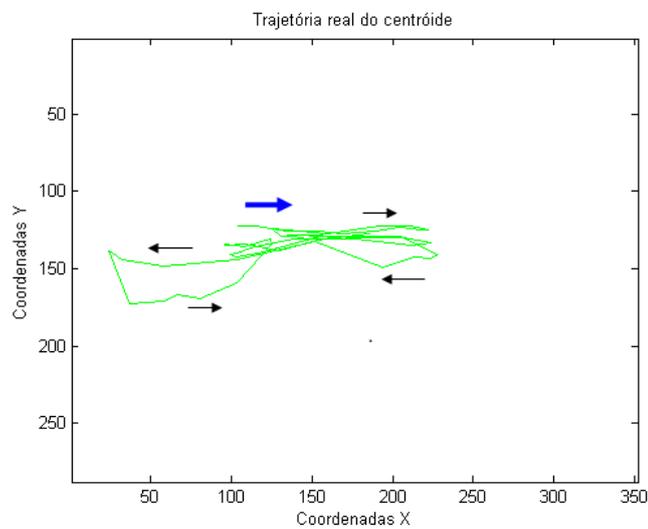


Figura 17 – Trajetória do centro do objeto alvo no vídeo 4.

<sup>6</sup> As setas indicam o sentido do objeto alvo no vídeo, onde a seta em azul o sentido inicial.

## 5.2 METODOLOGIA DE COMPARAÇÃO DE DESEMPENHO

As medidas de desempenho de um rastreador foram baseadas em WU et al. (2013). São consideradas duas formas de avaliação: (i) acurácia e (ii) robustez. A acurácia é uma medida que fornece o quanto a região detectada pelo rastreador sobrepõe a região que representa o objeto alvo. A robustez mede quantas vezes o rastreador perde o objeto alvo durante um conjunto de quadros.

### 5.2.1 Acurácia

A acurácia de um rastreador em um quadro  $i$  é definida pela sobreposição entre a caixa delimitadora prevista pelo rastreador  $A_i^T$  e a caixa delimitadora da verdadeira localização (*ground-truth*) do objeto alvo  $A_i^G$ :

$$\phi_i = \frac{A_i^G \cap A_i^T}{A_i^G \cup A_i^T}$$

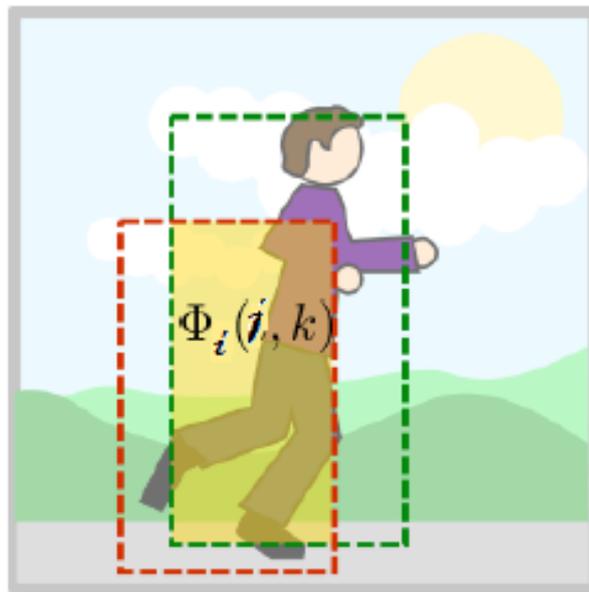


Figura 18 – Visualização da Acurácia. Adaptada de KRISTAN et al. (2013).

Considere que cada algoritmo executa quinze vezes cada vídeo,  $N_{rep} = 15$ , e que a acurácia do  $j$ -ésimo rastreador no quadro  $i$  da repetição  $k = 1, \dots, N_{rep}$  é  $\phi_i(j, k)$  e  $I_{valido}^k$  a

quantidade de quadros válidos na repetição  $k$ , ou seja, quadros livres da ação de reinicialização <sup>7</sup>. Logo, a acurácia média de um rastreador  $j$  na repetição  $k$  é dada por:

$$\phi(j, k) = \frac{1}{I_{\text{valido}}^k} \sum_{i=1}^{I_{\text{valido}}^k} \phi_i(j, k).$$

A acurácia média de um rastreador  $j$  é dada por:

$$\rho A(j) = \frac{1}{N_{\text{rep}}} \sum_{k=1}^{N_{\text{rep}}} \phi(j, k).$$

### 5.2.2 Robustez

Neste trabalho adota-se como perda de rastreamento quando a acurácia iguala a zero. Neste caso, o rastreamento é reinicializado manualmente.

A reinicialização não é feita imediatamente no quadro seguinte. Antes de iniciar a medida de acurácia é necessário desconsiderar dois grupos de quadros:  $\Delta_f$  e  $\Delta_o$ .

Em diversas situações o rastreador perde o objeto alvo e provavelmente continuará falhando nos quadros seguintes. Para diminuir essa correlação, o rastreador é reinicializado 5 quadros após a falha de predição ( $\Delta_f = 5$ ) (FERNANDEZ; VOJIR, 2013). A Figura 19 ilustra esse processo:

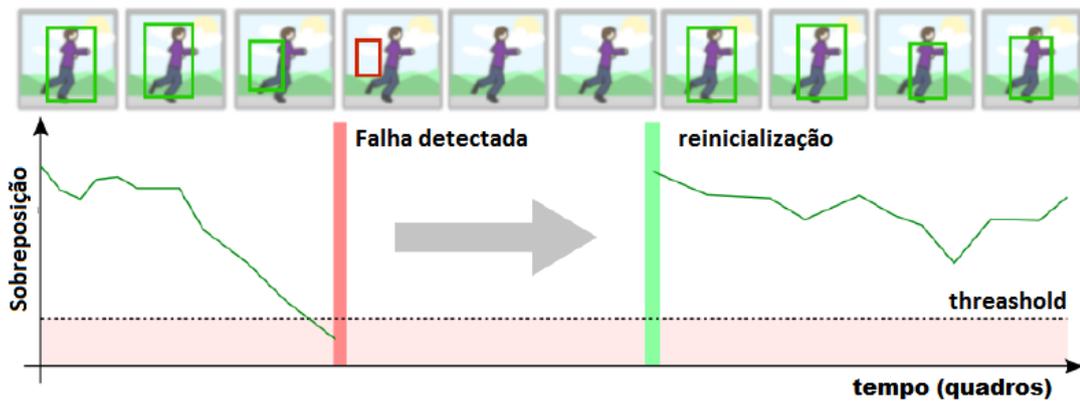


Figura 19 – Visualização da robustez. Adaptada de KRISTAN et al. (2013).

A acurácia dos quadros seguintes à reinicialização possuem valores tendenciosos, com máxima sobreposição para inicialização não-ruídosa, necessitando de alguns quadros para

<sup>7</sup> A reinicialização apenas acontece quando o valor de acurácia é menor que um *threshold*, que nos experimentos foi definido sendo igual a zero.

que esse valor se estabilize. Nos experimentos foram definidos que são necessários 10 quadros ( $\Delta_O = 10$ ) para que se estabilize, determinado empiricamente (FERNANDEZ; VOJIR, 2013).

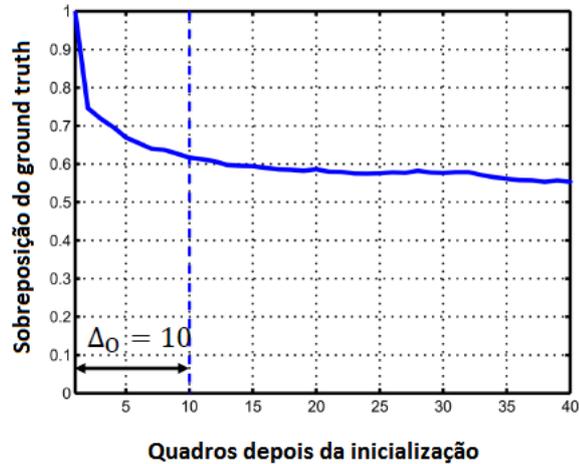


Figura 20 – Bias de Reinicialização. Adaptada de FERNANDEZ; VOJIR (2013).

Logo, os quadros  $\Delta_F$  e  $\Delta_O$  não são considerados válidos.

Considerando  $\gamma(j, k)$  o número de falhas do  $j$ -ésimo rastreador na repetição  $k$ , a robustez média de um rastreador  $j$  é definida como:

$$\rho R(j) = \frac{1}{N_{rep}} \sum_{k=1}^{N_{rep}} \gamma(j, k)$$

Para a comparação direta entre os vídeos, a medida de *robustez média absoluta*  $\tau R(j)$  é dada por:

$$\tau R(j) = \frac{\rho R(j)}{I_{total}}$$

sendo  $I_{total}$  o número de quadros do vídeo.

### 5.3 CONCLUSÕES DO CAPÍTULO

Neste capítulo foram mostrados os vídeos utilizados para testar os rastreadores e como foi feita a rotulação dos vídeos. Na seção 5.2 foram apresentadas as metodologias de comparação de cada rastreador utilizada neste trabalho, que são: acurácia e robustez. No próximo capítulo vão ser mostrados os resultados encontrados para cada rastreador.

## 6 RESULTADOS E DISCUSSÃO

Os resultados experimentais obtidos para as técnicas de rastreamento com Casamento de *Template*, Filtro de Kalman e Filtro de Partículas são expressos qualitativamente pela inspeção visual dos resultados do rastreamento e, quantitativamente, através da comparação estatística dos valores encontrados para a acurácia e robustez.

### 6.1 RESULTADOS

Cada rastreador foi executado quinze vezes,  $N_{rep} = 15$ , para cada vídeo, onde nas execuções a seleção do objeto alvo teve pequenas variações de tamanho e localização, com perturbação de  $\pm 20\%$  do tamanho real da região do objeto alvo. Para cada vídeo os resultados dos rastreadores são expressos em uma tabela, mostrando o valor da média da acurácia  $\phi(j, k)$ , valor da robustez  $\gamma(j, k)$  e o número do quadro que ocorreu a primeira perda (PP) total do objeto alvo pelo rastreador em cada execução.

São mostrados na Tabela 16 a média de quadros válidos de cada sequência de quadros analisados. Por fim é mostrada num **Mapa A-R** (Acurácia-Robustez) a média dos valores encontrados das métricas acurácia e robustez nas 15 execuções de cada rastreador por vídeo, e também, num **Mapa A-R** os valores médios das métricas acurácia e robustez de cada rastreador sobre todos os vídeos.

#### 6.1.1 Vídeo 1

Os resultados dos rastreadores para o vídeo 1 estão apresentados na Tabela 8.

Observe que acurácia é um índice do tipo “quanto maior, melhor” enquanto a robustez é do tipo “quanto menor, melhor”.

Os testes de hipótese CT-FK, CT-FP e FK-FP com a hipótese nula de que as médias são iguais para acurácia e robustez são expressões na Tabela 9.

Tabela 8 – Resultados do vídeo 1.

Execução	CT			FK			FP		
	$\phi(j,k)$	$\gamma(j,k)$	PP	$\phi(j,k)$	$\gamma(j,k)$	PP	$\phi(j,k)$	$\gamma(j,k)$	PP
01	0,111	10	138	0,457	2	156	0,464	2	156
02	<b>0,316</b>	<b>0</b>	-	0,300	6	133	0,211	3	155
03	0,184	9	173	0,402	2	148	0,235	2	155
04	<b>0,609</b>	<b>0</b>	-	0,229	4	130	0,078	4	149
05	0,115	12	140	0,240	5	129	0,272	2	155
06	0,202	7	176	0,651	2	155	0,312	2	155
07	0,118	10	171	0,180	5	146	0,207	3	156
08	0,207	9	173	0,467	3	129	0,342	2	155
09	0,196	9	138	<b>0,723</b>	3	155	0,380	2	155
10	0,218	9	140	0,518	2	148	0,406	2	155
11	<b>0,661</b>	<b>0</b>	-	0,490	2	148	0,300	2	155
12	0,161	8	139	0,142	5	130	0,327	2	154
13	0,138	9	138	0,258	3	152	0,323	2	155
14	0,236	9	131	0,627	2	165	0,464	3	154
15	0,285	8	138	0,600	2	156	0,366	2	155
Média	0,250	7,266		<b>0,419</b>	3,2		0,312	<b>2,333</b>	
Desvio padrão	0,167	3,918		0,186	1,424		0,103	0,617	

Tabela 9 – Teste de hipótese - vídeo 1.

	CT-FK	CT-FP	FK=FP
<i>p-value</i> acurácia	0,019	0,318	0,015
<i>p-value</i> robustez	0,003	0,001	0,022

Levando em consideração a significância de 5% os resultados mostram que para CT-FK é rejeitada a hipótese nula de que as médias são iguais tanto da acurácia quanto da robustez, ou seja, é considerado que as médias são diferentes. O mesmo ocorre para FK-FP nas médias da acurácia e robustez e no CT-FP para a média da robustez. Já no caso CT-FP não se pode afirmar que há diferença entre as médias da acurácia.

Para o vídeo 1 o resultado com o Filtro de Kalman possui a melhor média da acurácia e também obtendo, na nona execução, o melhor valor de acurácia. No entanto, o rastreador com Filtro de Partículas foi o que obteve a melhor média de valor de robustez. O rastreador com Casamento de *Template* obteve a pior média de acurácia e foi o que mais reinicializou, porém, dependendo da representação selecionada para o objeto alvo e sua localização, obteve os melhores resultados, por exemplo, os da execução 02, 04 e 11, não necessitando ser reinicializado.

Na Figura 21 é mostrado a média dos valores da acurácia encontrados nas quinze execuções de cada rastreador.

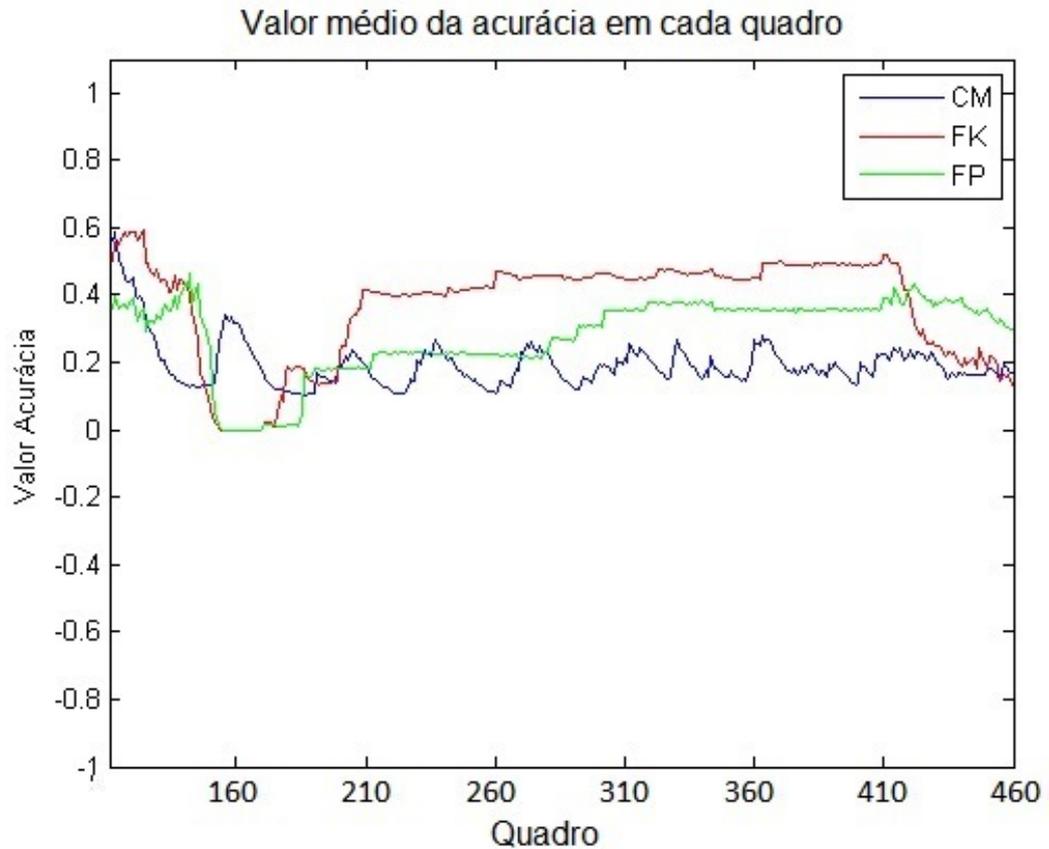


Figura 21 – Média dos valores da acurácia das quinze execuções para cada quadro do vídeo 1.

Na Figura 21 é possível identificar que em torno do quadro 140 ao quadro 190 e nos últimos 40 quadros, quando o objeto fazia manobras, os desempenhos dos rastreadores com Filtro de Kalman e Filtro de Partículas pioram. Em torno do quadro 190 ao quadro 420 o objeto alvo permanece parado, no entanto, vem mudando de aparência, fazendo com que os rastreadores com Filtro de Kalman e Filtro de Partículas obtenha melhores resultados do que o Casamento de *Template*, porém em alguns casos o Filtro de Partículas apenas não obtinha resultados melhores do que o Casamento de *Template* quando o Casamento de *Template* era reinicializado.

Alguns quadros com o rastreamento feito com Casamento de *Template*<sup>1</sup>, Filtro de Kalman<sup>2</sup> e Filtro de Partículas<sup>3</sup> criados para uma análise qualitativa são:

<sup>1</sup> O retângulo de cor verde representa a observação.

<sup>2</sup> O retângulo de cor preta representa a observação.

<sup>3</sup> O retângulo de cor vermelha representa a observação.



Figura 22 – Imagem com os quadros do vídeo 1 para o rastreador com Casamento de *Template*.

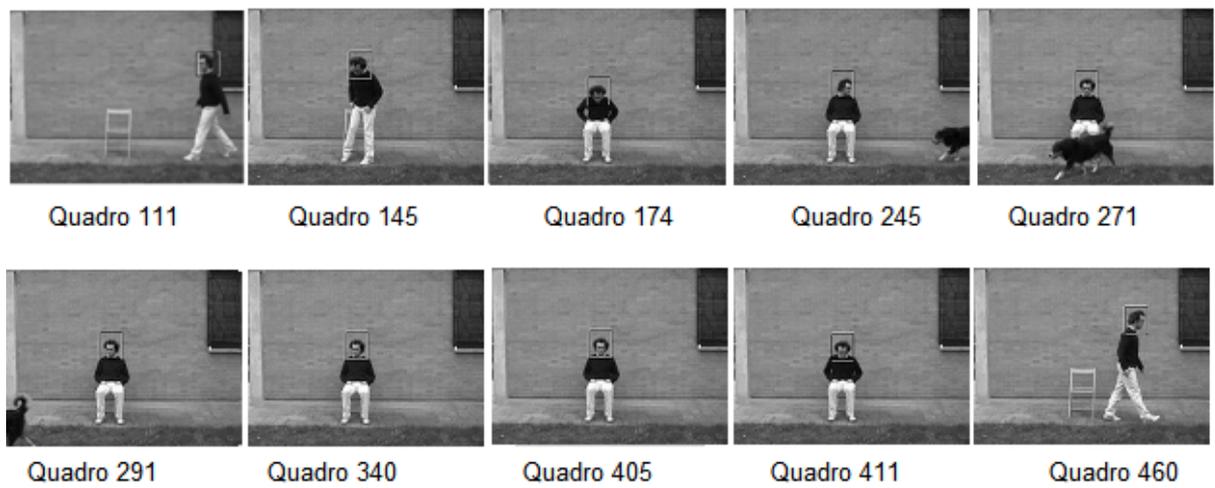


Figura 23 – Imagem com os quadros do vídeo 1 para o rastreador com Filtro de Kalman.

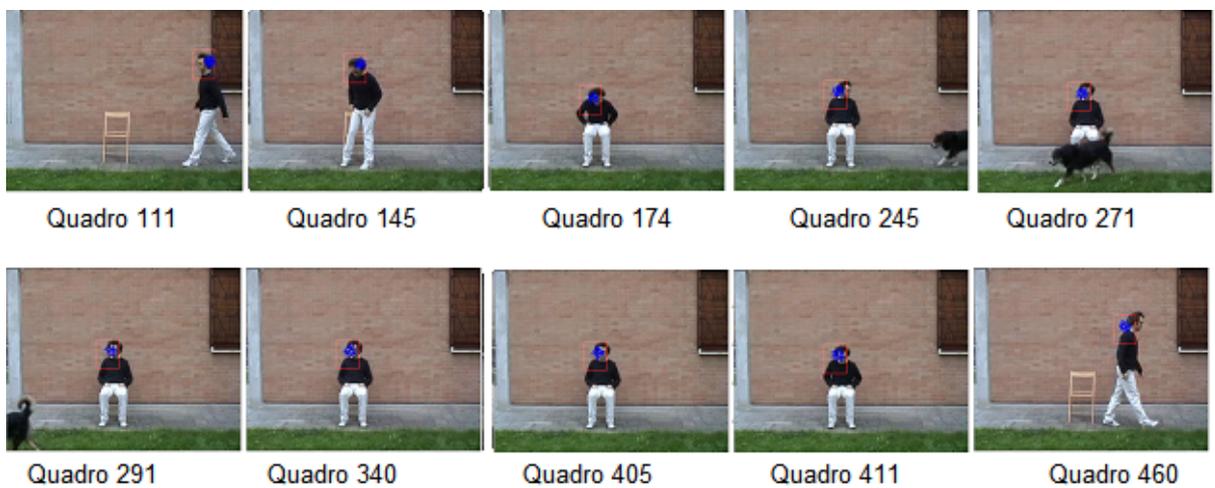


Figura 24 – Imagem com os quadros do vídeo 1 para o rastreador com Filtro de Partículas.

### 6.1.2 Vídeo 2

Os resultados dos rastreadores para o vídeo 2 estão apresentados na Tabela 10.

Tabela 10 – Resultados do vídeo 2.

Execução	CT			FK			FP		
	$\phi(j,k)$	$\gamma(j,k)$	PP	$\phi(j,k)$	$\gamma(j,k)$	PP	$\phi(j,k)$	$\gamma(j,k)$	PP
01	0,235	16	135	0,385	8	270	0,449	4	269
02	0,216	<b>2</b>	359	0,113	15	184	0,160	13	191
03	0,153	18	135	0,521	4	335	<b>0,595</b>	6	273
04	0,170	16	133	0,277	5	256	0,311	7	257
05	0,124	20	132	0,145	17	195	0,172	10	196
06	0,152	16	140	0,440	6	254	0,438	8	260
07	0,197	15	146	0,361	9	265	0,441	5	271
08	0,143	19	135	0,170	14	197	0,204	12	196
09	0,198	16	142	0,459	10	223	0,345	7	231
10	0,273	15	137	0,371	4	270	0,485	5	320
11	0,136	20	133	0,242	11	199	0,301	10	205
12	0,324	3	<b>362</b>	0,172	13	199	0,212	11	197
13	0,248	14	142	0,342	7	263	0,394	4	269
14	0,150	17	132	0,134	12	194	0,210	12	227
15	0,415	4	<b>362</b>	0,297	8	238	0,287	7	258
Média	0,209	14,067		0,295	9,533		<b>0,334</b>	<b>8,067</b>	
Desvio padrão	0,080	6,006		0,129	4,068		0,130	3,058	

Os testes de hipótese CT-FK, CT-FP e FK-FP com a hipótese nula de que as médias são iguais para acurácia e robustez são expressões na Tabela 11.

Tabela 11 – Teste de hipótese - vídeo 2.

	CT-FK	CT-FP	FK=FP
<i>p-value</i> acurácia	0,042	0,007	0,013
<i>p-value</i> robustez	0,040	0,007	0,045

Levando em consideração a significância de 5% os resultados mostram que para CT-FK, CT-FP e FK-FP é rejeitada a hipótese nula de que as médias são iguais tanto da acurácia quanto da robustez, ou seja, é considerado que as médias são diferentes.

Para o vídeo 2 o rastreador com Filtro de Partículas foi o que obteve a melhor média de acurácia e a melhor média de robustez e também, na terceira execução o maior valor de

acurácia. O rastreador com Casamento de *Template* foi o que demorou mais a perder por total o objeto alvo (nas execuções 12 e 15 rastreou até o quadro 362) e o que obteve, na execução 02, a menor robustez.

Na Figura 25 é mostrado a média dos valores da acurácia encontrados nas quinze execuções de cada rastreador.

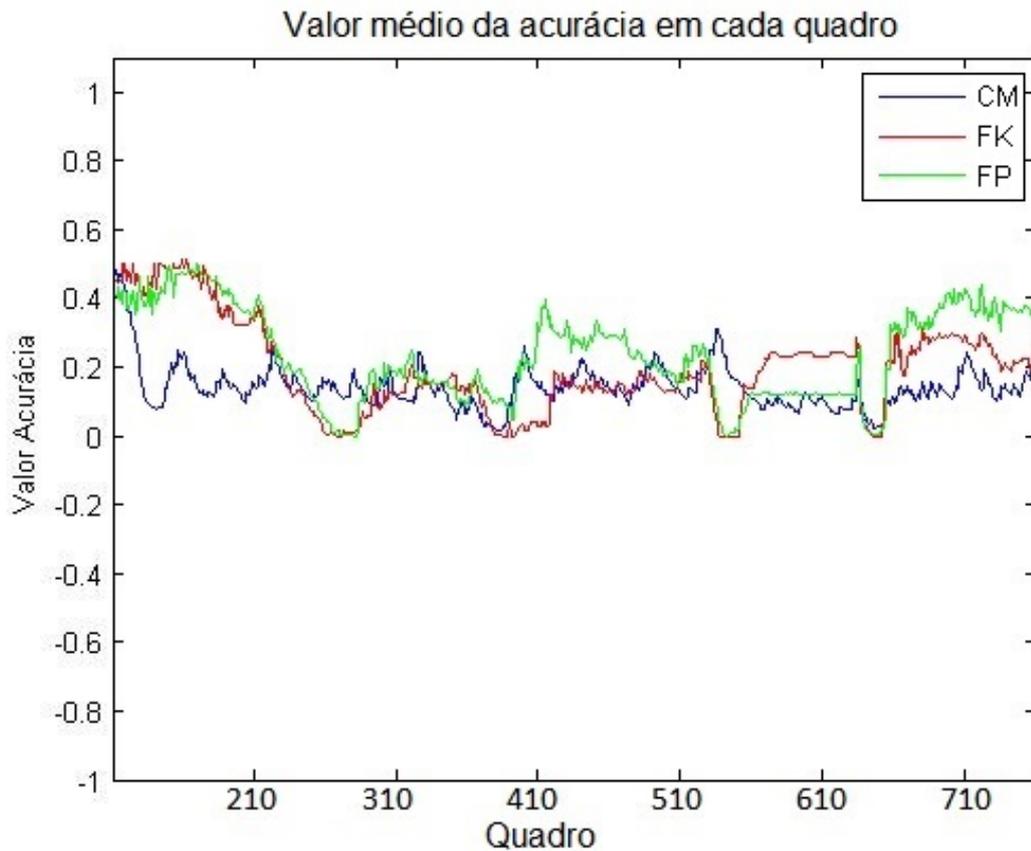


Figura 25 – Média dos valores da acurácia das quinze execuções para cada quadro do vídeo 2.

A Figura 25 mostra que as médias dos rastreadores são bem parecidas, o Filtro de Kalman e Filtro de Partículas pioram o seu desempenho em torno do quadro 210 ao quadro 270 devido o objeto alvo realizar manobras. Em torno do quadro 400 todos os rastreadores pioram o desempenho devido o objeto alvo passar por oclusões.

Alguns quadros com o rastreamento feito com Casamento de *Template*<sup>4</sup>, Filtro de Kalman<sup>5</sup> e Filtro de Partículas<sup>6</sup> criados para uma análise qualitativa são:

<sup>4</sup> O retângulo de cor verde representa a observação.

<sup>5</sup> O retângulo de cor preta representa a observação.

<sup>6</sup> O retângulo de cor vermelha representa a observação.

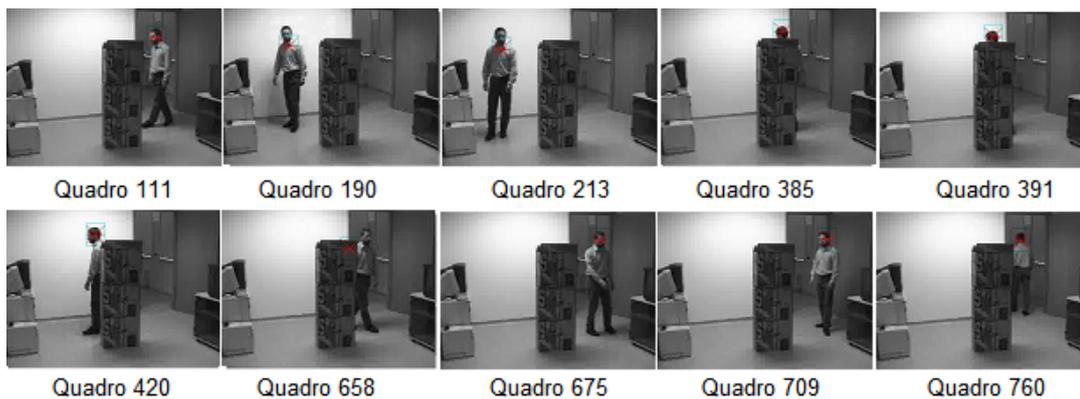


Figura 26 – Imagem com os quadros do vídeo 2 para o rastreador com Casamento de *Template*.

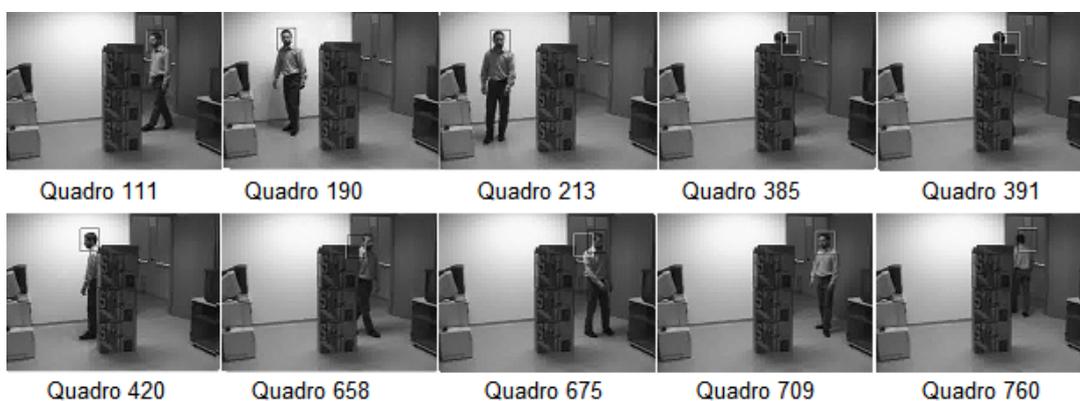


Figura 27 – Imagem com os quadros do vídeo 2 para o rastreador com Filtro de Kalman.

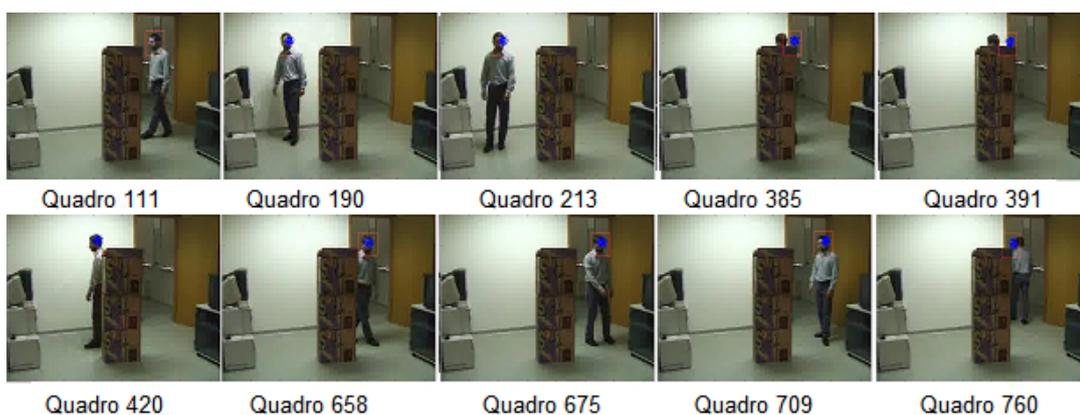


Figura 28 – Imagem com os quadros do vídeo 2 para o rastreador com Filtro de Partículas.

### 6.1.3 Vídeo 3

Os resultados dos rastreadores para o vídeo 3 estão apresentados na Tabela 12.

Tabela 12 – Resultados do vídeo 3.

Execução	CT			FK			FP		
	$\phi(j,k)$	$\gamma(j,k)$	PP	$\phi(j,k)$	$\gamma(j,k)$	PP	$\phi(j,k)$	$\gamma(j,k)$	PP
01	0,236	17	134	0,314	18	127	0,315	15	129
02	0,435	10	139	0,277	19	133	0,338	14	129
03	0,135	19	162	0,159	18	122	0,184	17	133
04	0,254	15	132	0,360	12	126	<b>0,553</b>	13	128
05	0,174	18	152	0,181	21	123	0,226	19	127
06	0,177	18	164	0,305	22	133	0,252	17	132
07	0,132	17	137	0,294	19	127	0,286	18	129
08	0,281	16	135	0,307	15	128	0,258	17	126
09	0,388	<b>8</b>	156	0,229	21	133	0,271	16	129
10	0,277	15	134	0,312	19	133	0,280	17	133
11	0,326	17	134	0,266	18	127	0,286	18	132
12	0,193	9	152	0,124	20	131	0,158	19	135
13	0,196	15	<b>168</b>	0,200	21	126	0,388	15	128
14	0,445	11	139	0,302	18	127	0,314	14	129
15	0,199	16	135	0,210	21	133	0,297	18	127
Média	0,256	<b>14,733</b>		0,256	18,800		<b>0,294</b>	16,467	
Desvio padrão	0,102	3,515		0,068	2,597		0,092	1,885	

Os testes de hipótese CT-FK, CT-FP e FK-FP com a hipótese nula de que as médias são iguais para acurácia e robustez são expressões na Tabela 13.

Tabela 13 – Teste de hipótese - vídeo 3.

	CT-FK	CT-FP	FK=FP
<i>p-value</i> acurácia	0,987	0,247	0,066
<i>p-value</i> robustez	0,004	0,074	0,002

Levando em consideração a significância de 5% os resultados mostram que para CT-FK, CT-FP e FK-FP não é rejeitada a hipótese nula de que as médias são iguais para a acurácia, ou seja, não pode ser afirmado que as médias são diferentes. Também não pode ser afirmado que a média é diferente da robustez para CT-FP, no entanto para CT-FK e FK-FP é rejeitada a hipótese nula de que as médias são iguais para a robustez.

Devido as dificuldades existentes no vídeo 3, tais como manobras bruscas, oclusões, mudanças de aparência e até mudança de escala do objeto alvo, nenhum método de rastreamento obteve resultados satisfatório em termos de robustez. Entre os três rastreadores, o que obteve a melhor média de robustez foi o Casamento de *Template*, no entanto, a melhor média de acurácia foi o rastreador com Filtro de *Partículas*. Um dos fatores dos resultados obtidos pelo Filtro de

Kalman e Filtro de Partículas, em termos de robustez, foi devido aos movimentos bruscos do objeto alvo, fazendo com que o modelo dinâmico adotado para eles não fosse o mais adequado. O rastreador com Casamento de *Template* conseguiu rastrear por mais tempo, na execução 13 rastreou até o quadro **168**, devido ao fato de que, apesar das manobras, o objeto alvo não sofreu grandes mudanças de aparência.

Na Figura 29 é mostrado a média dos valores da acurácia encontrados nas quinze execuções de cada rastreador.

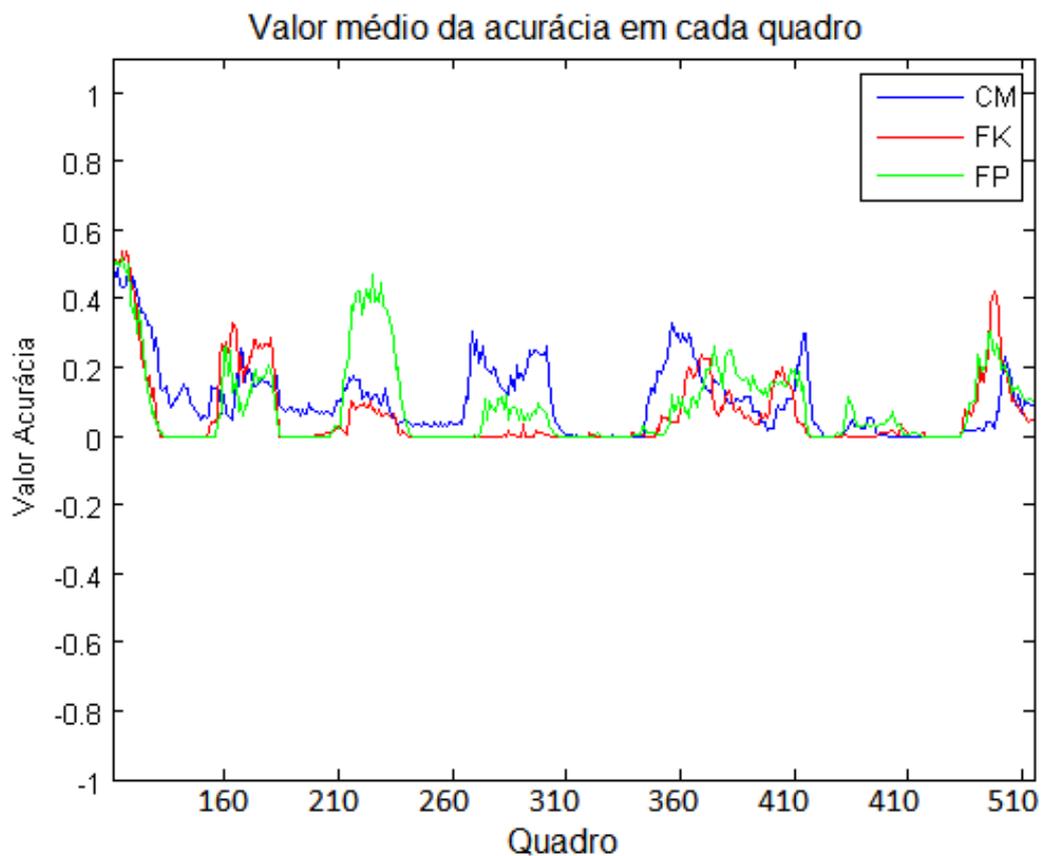


Figura 29 – Média dos valores da acurácia das quinze execuções para cada quadro do vídeo 3.

Na Figura 29 mostra que as médias dos rastreadores são bem parecidas, em vários quadros os rastreadores obtiveram médias baixas (próximas de zero) devido o objeto alvo passar por logos períodos de oclusão total, pois os rastreadores perdiam o objeto alvo e quando era reinicializado a acurácia de quinze quadros após a inicialização eram considerados iguais a zero.

Alguns quadros com o rastreamento feito com Casamento de *Template*<sup>7</sup>, Filtro de Kalman<sup>8</sup> e Filtro de Partículas<sup>9</sup> criados para uma análise qualitativa são:

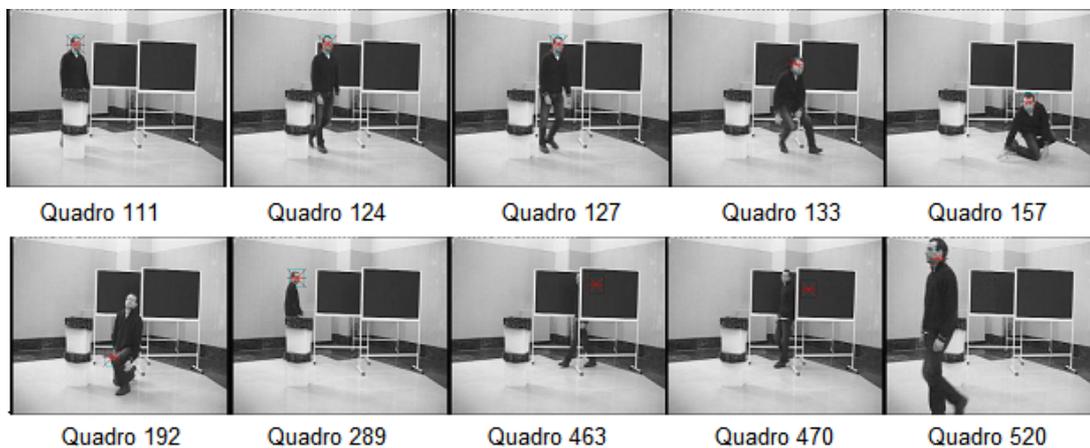


Figura 30 – Imagem com os quadros do vídeo 3 para o rastreador com Casamento de *Template*.

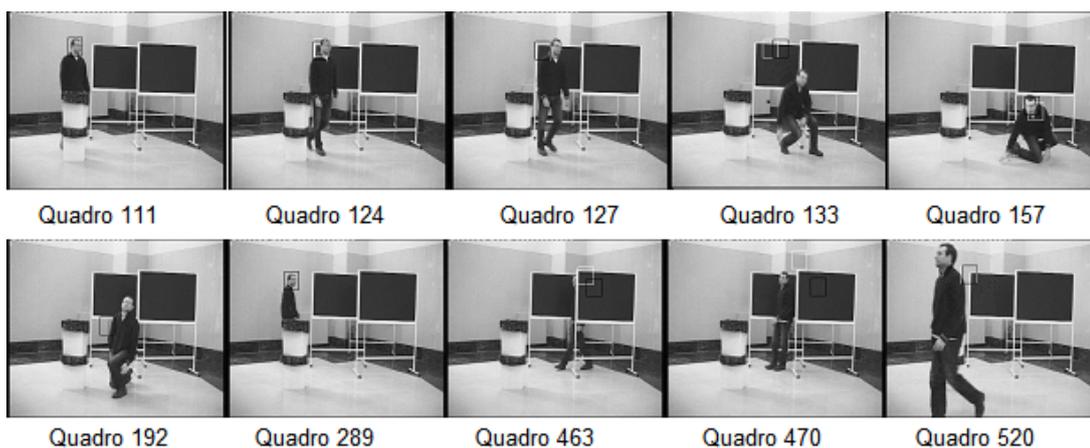


Figura 31 – Imagem com os quadros do vídeo 3 para o rastreador com Filtro de Kalman.

<sup>7</sup> O retângulo de cor verde representa a observação.

<sup>8</sup> O retângulo de cor preta representa a observação.

<sup>9</sup> O retângulo de cor vermelha representa a observação.

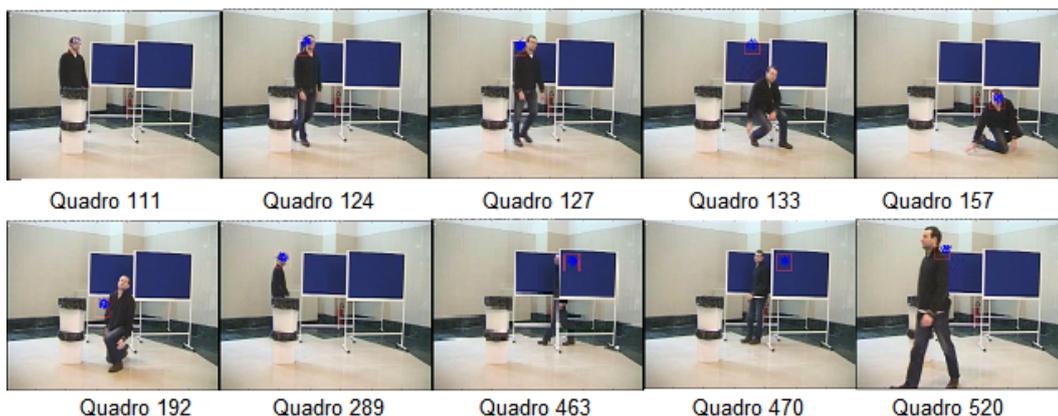


Figura 32 – Imagem com os quadros do vídeo 3 para o rastreador com Filtro de Partículas.

#### 6.1.4 Vídeo 4

Os resultados dos rastreadores para o vídeo 4 estão apresentados na Tabela 14.

Tabela 14 – Resultados do vídeo 4.

Execução	CT			FK			FP		
	$\phi(j,k)$	$\gamma(j,k)$	PP	$\phi(j,k)$	$\gamma(j,k)$	PP	$\phi(j,k)$	$\gamma(j,k)$	PP
01	0,382	9	13	0,365	7	40	0,474	6	38
02	0,344	9	12	0,251	7	40	0,339	6	37
03	0,152	8	14	0,467	6	32	0,426	5	37
04	0,025	8	15	0,440	7	16	0,329	<b>4</b>	<b>66</b>
05	0,075	9	12	0,497	6	14	0,477	6	38
06	0,063	9	12	0,440	6	34	0,511	5	36
07	0,361	9	12	0,277	6	38	0,299	7	35
08	0,183	10	12	0,309	6	39	0,411	5	36
09	0,144	7	14	0,483	6	14	0,415	6	38
10	0,121	9	13	0,424	6	33	0,462	5	39
11	0,259	9	11	0,290	5	37	0,287	6	34
12	0,334	9	12	0,388	7	32	0,403	6	36
13	0,294	9	12	0,331	7	35	0,456	6	36
14	0,178	7	14	0,434	6	32	0,435	5	35
15	0,177	7	14	0,493	7	17	<b>0,516</b>	6	36
Média	0,206	8,533		0,392	6,333		<b>0,416</b>	<b>5,600</b>	
Desvio padrão	0,115	0,915		0,083	0,617		0,073	0,737	

Os testes de hipótese CT-FK, CT-FP e FK-FP com a hipótese nula de que as médias são iguais para acurácia e robustez são expressões na Tabela 15.

Tabela 15 – Teste de hipótese - vídeo 4.

	CT-FK	CT-FP	FK=FP
<i>p-value</i> acurácia	0,002	0,0001	0,198
<i>p-value</i> robustez	0,000003	0,0000001	0,010

Levando em consideração a significância de 5% os resultados mostram que para CT-FK, CT-FP e FK-FP é rejeitada a hipótese nula de que as médias são iguais para a robustez, ou seja, pode ser afirmado que as médias são diferentes. Também pode ser afirmado que a média é diferente da acurácia para CT-FK e CT-FP, no entanto para FK-FP não é rejeitada a hipótese nula de que as médias são iguais para a acurácia.

No vídeo 4 o rastreador com Filtro de Partículas foi o que obteve os melhores resultados. Os rastreadores com Filtro de Kalman e Filtro de Partículas não encontraram muitas dificuldades em rastrear o objeto alvo quando o mesmo sofria mudanças de iluminação, no entanto, apresentou dificuldades ao rastrear com as constantes mudanças de direção. O rastreador com Casamento de *Template* perdeu o objeto alvo logo que o mesmo sofreu a mudança de iluminação, tendo, por isso, que ser reinicializado logo nos primeiros quadros.

Na Figura 33 é mostrado a média dos valores da acurácia encontrados nas quinze execuções de cada rastreador.

Na Figura 33 mostra que o rastreador Casamento de *Template* obteve na maioria dos quadros média de acurácia igual a zero, pois o objeto alvo passava por constantes mudança de iluminação. Os rastreadores com Filtro de Kalman e Filtro de Partículas também obteve médias, em vários quadros, igual a zero devido as constantes manobras do objeto alvo. O rastreador com Casamento de *Template* apenas possuiu médias diferentes de zero quando o objeto alvo após quinze quadros depois de uma inicialização não passava por mudança de iluminação. Os rastreadores com Filtro de Kalman e Filtro de Partículas melhoravam os seus desempenhos quando o objeto alvo após quinze quadros não passava por manobras.

Alguns quadros com o rastreamento feito com Casamento de *Template*<sup>10</sup>, Filtro de Kalman<sup>11</sup> e Filtro de Partículas<sup>12</sup> criados para uma análise qualitativa são:

<sup>10</sup> O retângulo de cor verde representa a observação.

<sup>11</sup> O retângulo de cor preta representa a observação.

<sup>12</sup> O retângulo de cor vermelha representa a observação.

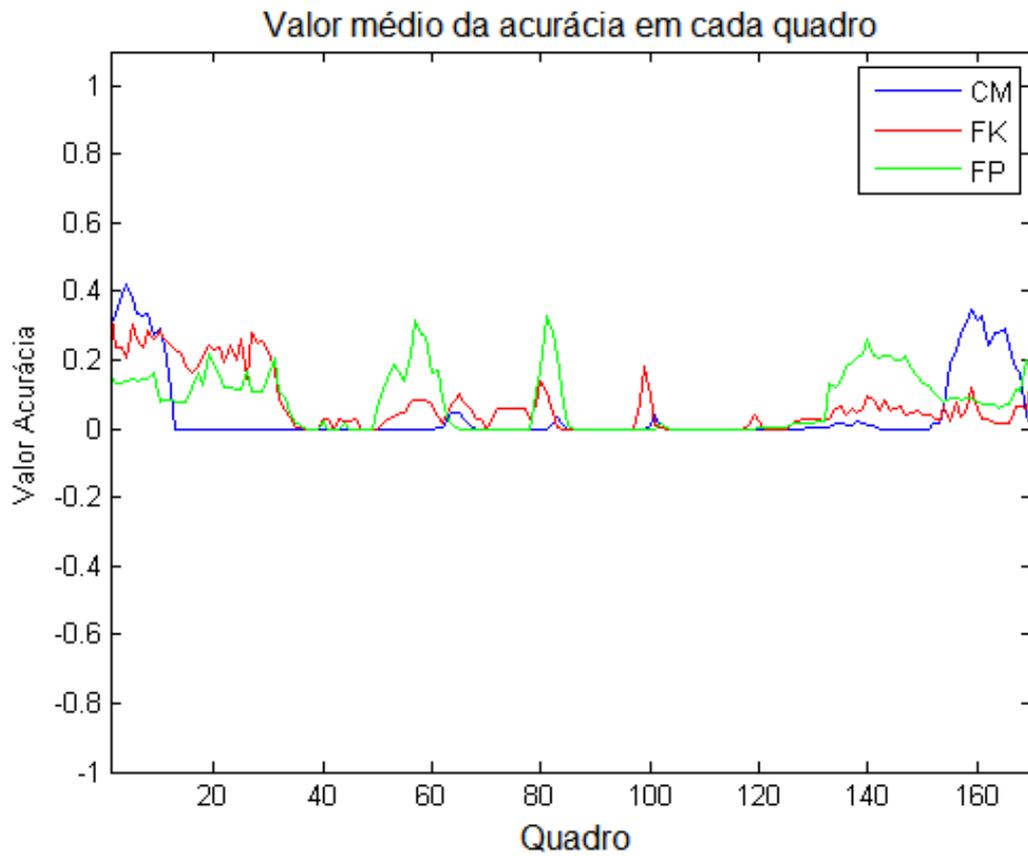


Figura 33 – Média dos valores da acurácia das quinze execuções para cada quadro do vídeo 4.



Figura 34 – Imagem com os quadros do vídeo 4 para o rastreador com Casamento de *Template*.



Figura 35 – Imagem com os quadros do vídeo 4 para o rastreador com Filtro de Kalman.



Figura 36 – Imagem com os quadros do vídeo 4 para o rastreador com Filtro de Partículas.

## 6.2 DISCUSSÃO

O rastreador com Filtro de Partículas obteve o melhor desempenho, com a média da acurácia igual a 0,338 e robustez igual a 8,116, já as médias da acurácia e robustez do Filtro de Kalman, respectivamente, são 0,340 e 9,466, e os resultados do Casamento de *Template* são 0,230 para acurácia e 11,150 para robustez. Na tabela 16, que contém a média de quadros válidos de cada vídeo, mostra que o rastreador com a técnica Filtro de Partículas obteve a melhor média.

Os resultados quantitativos de cada rastreador por vídeo, levando em consideração a acurácia média e a robustez média absoluta, e os resultados dos rastreadores, levando em

Tabela 16 – Média de quadros válidos dos algoritmos por vídeo.

Vídeos	CT	FK	FP
Vídeo 1	242 (69%)	302 (86%)	315 (90%)
Vídeo 2	440 (68%)	508 (78%)	530 (82%)
Vídeo 3	190 (46%)	128 (31%)	164 (40%)
Vídeo 4	46 (27%)	76 (45%)	87 (51%)
Média	52,5%	60%	<b>65,7%</b>

consideração a média da acurácia e média da robustez de todas as execuções, podem ser visto nas figuras 37 e 38, respectivamente.

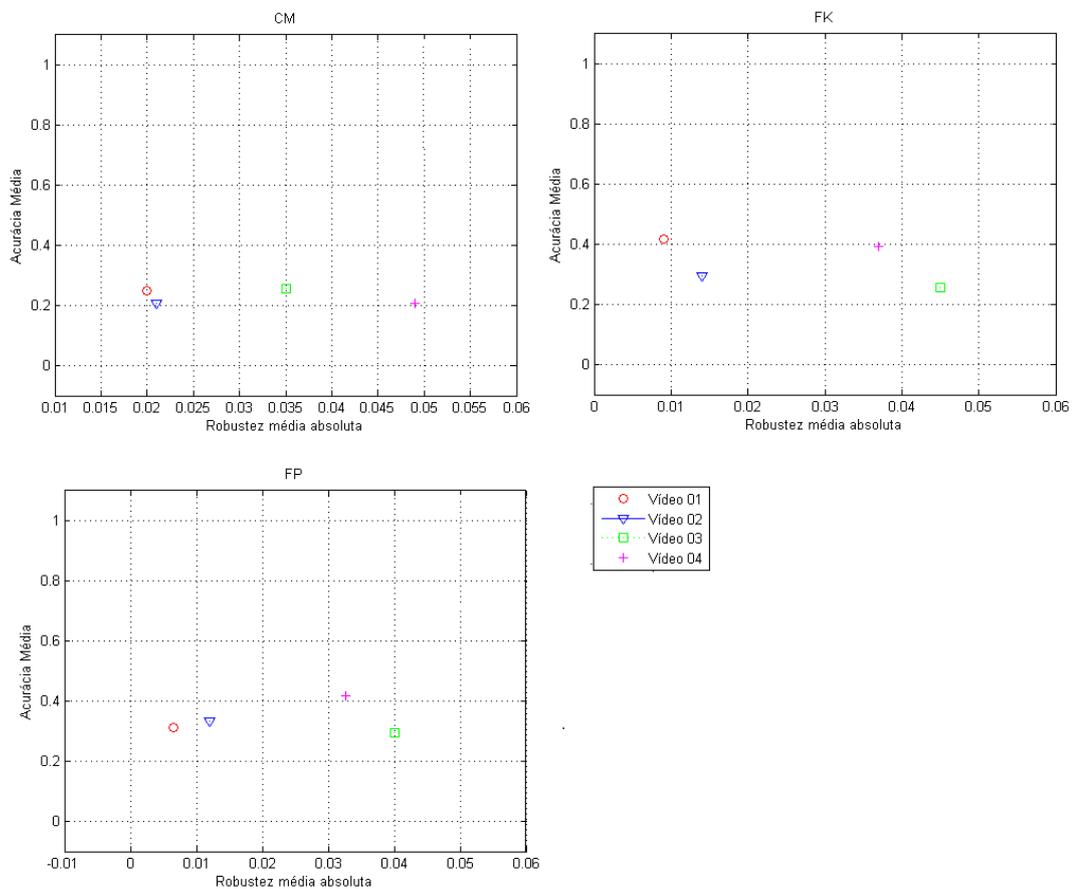


Figura 37 – Gráficos A-R de cada rastreador por vídeo.

Analisando o gráfico anterior, onde indica que quanto mais próximo os pontos se encontram do canto superior esquerdo, melhor é o desempenho do rastreador, percebe-se que o rastreador com Filtro de Partículas possui o melhor desempenho. Apesar dos resultados obtidos pelos rastreadores serem diferentes, existem situações onde cada rastreador teve o melhor desempenho, demonstrando que cada um possui características únicas que os beneficiam em

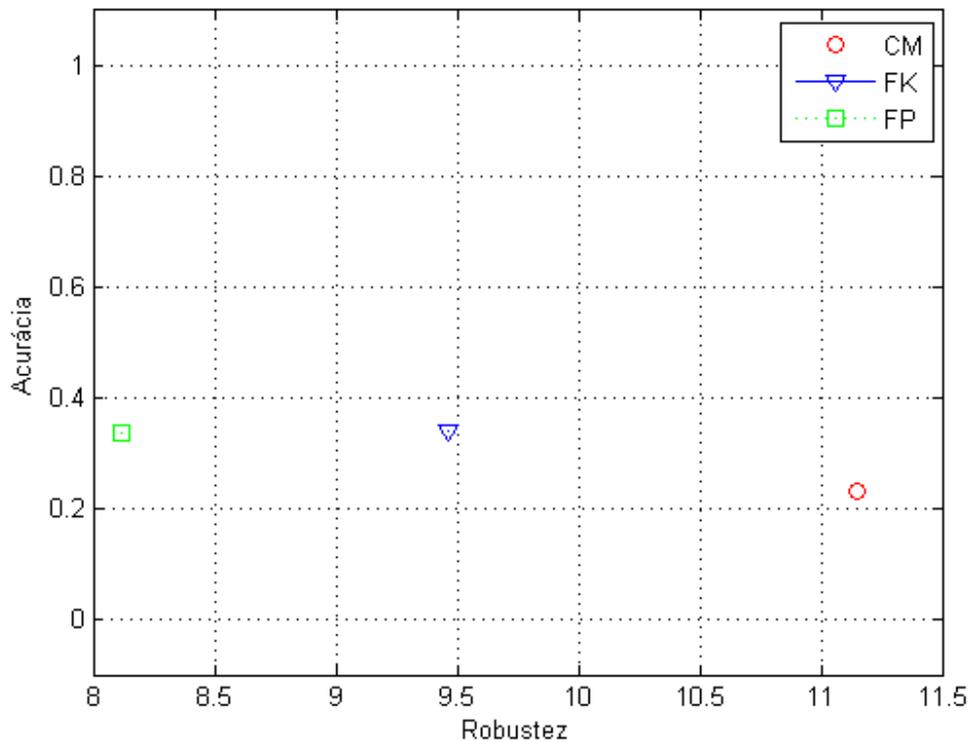


Figura 38 – Resultado dos rastreadores MAPA A-R.

algumas situações.

Os resultados dos rastreadores com relação as perturbações que o objeto alvo pode sofrer são:

Casamento de *Template* não rastreou na maioria dos testes quando o objeto alvo passava por mudanças de aparência total e nem quando sofria mudança de iluminação parcial, pelo fato das variações de valores dos pixels da região do objeto alvo de um quadro para outro, pois o rastreador com Casamento de *Template* tentava encontrar a região do objeto alvo que era representado pelo quadro anterior, no entanto, no quadro atual já era bem diferente, fazendo com que outras regiões fossem definidas como a região do objeto alvo.

Filtro de Kalman e Filtro de Partículas tiveram dificuldades em rastrear o objeto alvo quando o mesmo sofria manobra total. Como os rastreadores com Filtro de Kalman e Filtro de Partículas utilizam modelo dinâmico, quando o objeto alvo sofria manobra total, o modelo adotado não representava a manobra, fazendo com que o rastreador continuasse a detectar a região de acordo com o modelo adotado.

Todos os rastreadores não conseguiram rastrear o objeto alvo quando o mesmo passava por longo período de oclusão total, devido ao fato da representação do objeto alvo, após

Tabela 17 – Resumo dos resultados dos rastreadores.

	<b>Nenhuma Mudança de Aparência</b>	<b>Mudança de Aparência Parcial</b>	<b>Mudança de Aparência Total</b>
CT	Rastreou	Rastreou	Não rastreou
FK	Rastreou	Rastreou	Rastreou
FP	Rastreou	Rastreou	Rastreou
	<b>Nenhuma Oclusão</b>	<b>Oclusão Parcial</b>	<b>Oclusão Total</b>
CT	Rastreou	Rastreou	Não rastreou
FK	Rastreou	Rastreou	Não rastreou
FP	Rastreou	Rastreou	Não rastreou
	<b>Nenhuma Manobra</b>	<b>Manobra Parcial</b>	<b>Manobra Total</b>
CT	Rastreou	Rastreou	Rastreou
FK	Rastreou	Rastreou	Não rastreou
FP	Rastreou	Rastreou	Não rastreou
	<b>Nenhuma Mudança de Iluminação</b>	<b>Mudança de Iluminação Parcial</b>	<b>Mudança de Iluminação Total</b>
CT	Rastreou	Não rastreou	NÃO FOI TESTADO
FK	Rastreou	Rastreou	NÃO FOI TESTADO
FP	Rastreou	Rastreou	NÃO FOI TESTADO

longo período, não ser mais a verdadeira representação.

### 6.3 CONCLUSÕES DO CAPÍTULO

Neste capítulo foram apresentados e discutidos os resultados obtidos com as implementações dos rastreadores Casamento de *Template*, Filtro de Kalman e Filtro de Partículas. Foram aplicados experimentos para avaliar por uma métrica de acurácia e outra de robustez cada rastreador, além de mostrar até qual quadro o rastreador conseguiu rastrear.

Observando os resultados, percebe-se que o rastreador com Filtro de Partículas, de maneira geral, apresentou melhores resultados que os outros dois rastreadores, tendo os melhores valores de média da acurácia e média de robustez, conseguindo rastrear por mais quadros na maioria dos testes.

## 7 CONCLUSÕES E TRABALHOS FUTUROS

Neste trabalho foi realizada uma comparação entre três técnicas clássicas de rastreamento de objetos em vídeo: Casamento de *Template*, Filtro de Kalman e Filtro de Partículas. O trabalho está focado em determinar que técnica de rastreamento é mais adequada quando o objeto a ser rastreado está submetido a diferentes níveis de manobra, oclusão, iluminação e aparência.

Os testes mostraram que o rastreador com Filtro de Partículas apresentou melhores resultados, conseguindo rastrear por mais quadros na maioria dos testes do que o Filtro de Kalman e o Casamento de *Template*. Assim como o Filtro de Partículas, o Filtro de Kalman conseguiu rastrear quando o objeto alvo passava por diferentes níveis de mudanças de aparência, pequenas mudanças de iluminação e manobras suaves, porém, todas as técnicas testadas não conseguiram rastrear quando o objeto alvo passava por oclusões totais. O rastreador por Casamento de *Template* apenas possuía melhores resultados do que os outros rastreadores quando o objeto alvo passava por manobras totais.

Para trabalhos futuros, pode ser feita análise e testes para comparar o tempo de execução de cada rastreador estudado. Além disso, pode tentar outros modelos de dinâmica do objeto utilizados no Filtro de Kalman e Filtro de Partículas.

Como o desempenho do rastreamento por Filtro de Kalman e Filtro de Partículas está intimamente ligado ao desempenho da observação, análise e testes de outros mecanismos de observação se mostram interessantes para trabalhos futuros.

## REFERÊNCIAS

- AMARAL, W. M. do; TOZZI, C. L. Tracking people through occlusions using a head detection method. 2013.
- ARULAMPALAM, M. S.; MASKELL, S.; GORDON, N.; CLAPP, T. A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking. **Signal Processing, IEEE Transactions on**, IEEE, v. 50, n. 2, p. 174–188, 2002.
- BALLARD, D. H.; BROWN, C. M. Principles of animate vision. **CVGIP: Image Understanding**, Elsevier, v. 56, n. 1, p. 3–21, 1992.
- BEM, R. A. d. **Uma abordagem livre de modelo para rastreamento de objetos em seqüências de imagens**. Tese (Doutorado) — Universidade de São Paulo, 2007.
- BISHOP, G.; WELCH, G. An introduction to the kalman filter. **Proc of SIGGRAPH, Course**, v. 8, n. 27599-23175, p. 41, 2001.
- FERNANDEZ, G.; VOJIR, T. The vot2013 challenge; overview and additional results. 2013.
- FORSYTH, D. A.; PONCE, J. **Computer vision: a modern approach**. [S.l.]: Prentice Hall Professional Technical Reference, 2002.
- FRIEMEL, B. H.; BOHS, L. N.; TRAHEY, G. E. Relative performance of two-dimensional speckle-tracking techniques: normalized correlation, non-normalized correlation and sum-absolute-difference. In: IEEE. **Ultrasonics Symposium, 1995. Proceedings., 1995 IEEE**. [S.l.], 1995. v. 2, p. 1481–1484.
- GONZALES, R. C.; WOODS, R. E. **Digital image processing**. [S.l.]: Prentice hall Upper Saddle River, NJ:, 2002.
- GREWAL, M. S.; ANDREWS, A. P. **Kalman filtering: theory and practice using MATLAB**. [S.l.]: John Wiley & Sons, 2011.
- GRISSETTI, G.; STACHNISS, C. Ecmr tutorial learning grid maps with rao-blackwellized particle filters. 2007.
- ILHA, G. Rastreamento automático da bola de futebol em vídeos. 2009.
- KRISTAN, M.; PFLUGFELDER, R.; LEONARDIS, A.; MATAS, J.; PORIKLI, F.; CEHOVIN, L.; NEBEHAY, G.; FERNANDEZ, G.; VOJIR, T.; GATT, A. et al. The visual object tracking vot2013 challenge results. In: IEEE. **Computer Vision Workshops (ICCVW), 2013 IEEE International Conference on**. [S.l.], 2013. p. 98–111.
- NORVIG, P.; RUSSELL, S. **Inteligência Artificial, 2ª Edição**. [S.l.]: Elsevier Brasil, 2004. v. 1.
- NUMMIARO, K.; KOLLER-MEIER, E.; GOOL, L. V. An adaptive color-based particle filter. **Image and vision computing**, Elsevier, v. 21, n. 1, p. 99–110, 2003.
- O'HAGAN, R.; ZELINSKY, A. **Finger Track - A Robust and Real-Time Gesture Interface**. [S.l.]: Australian Joint Conference on Artificial Intelligence, Perth, 1997. 302 p.
- PFEIFFER, M. A brief introduction to particle filters. **documento pdf**, 2004.

RISTIC, B.; ARULAMPALAM, S.; GORDON, N. **Beyond the Kalman filter: Particle filters for tracking applications.** [S.l.]: Artech house Boston, 2004. v. 4.

SÄRKKÄ, S.; VEHTARI, A.; LAMPINEN, J. Rao-blackwellized monte carlo data association for multiple target tracking. In: **I. Proceedings of the seventh international conference on information fusion.** [S.l.], 2004. v. 1, p. 583–590.

WIDYAWAN, M. I. Z.; NUGROHO, L. E. Adaptive motion detection algorithm using frame differences and dynamic template matching method. 2012.

WU, Y.; LIM, J.; YANG, M.-H. Online object tracking: A benchmark. In: IEEE. **Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on.** [S.l.], 2013. p. 2411–2418.

YILMAZ, A.; JAVED, O.; SHAH, M. Object tracking: A survey. **Acm computing surveys (CSUR)**, ACM, v. 38, n. 4, p. 13, 2006.