

Roberto Franciscatto

**Classificação de grupos utilizando
informações de geometria e detecção de
atividades intragrupo**

São Leopoldo

2008

Roberto Franciscatto

**Classificação de grupos utilizando
informações de geometria e detecção de
atividades intragrupo**

Dissertação submetida a avaliação como re-
quisito parcial para a obtenção do grau de
Mestre em Computação Aplicada.

Orientador:
Claudio Rosito Jung

UNIVERSIDADE DO VALE DO RIO DOS SINOS
CIÊNCIAS EXATAS E TECNOLÓGICAS
PROGRAMA INTERDISCIPLINAR DE PÓS-GRADUAÇÃO EM
COMPUTAÇÃO APLICADA

São Leopoldo

2008

Agradecimentos

Agradeço primeiramente a Deus, razão de minha existência e profundo conhecedor dos meus atos.

Aos meus pais, que sempre foram meus exemplos de caráter, fé, dinamismo, perseverança e pelos sólidos conhecimentos de vida que ensinam a cada dia.

Não poderia deixar de agradecer aos meus colegas de laboratório e amigos da UNISINOS na qual convivi durante estes anos, em especial a pessoa de Júlio Jacques Jr., pelo incentivo, ajuda e por estar sempre colaborando e dando idéias para este trabalho.

Aos meus orientadores e professores Cláudio Rosito Jung e Soraia Raupp Musse, pelo exemplo de grandes profissionais, incentivadores e principalmente grandes pessoas humanitárias.

A minha esposa Laura, pela paciência, carinho, apoio e amor que demonstraste durante esta caminhada.

A HP do Brasil, pelo auxílio financeiro e apoio a pesquisa.

Resumo

Este trabalho tem como principal objetivo propor um modelo para classificação automática (ou semi-automática) de grupos de pessoas utilizando informações de geometria e detecção da atividade intragrupo em seqüências de vídeo. Para a classificação de grupos de pessoas, utiliza-se um algoritmo de acompanhamento de objetos para rastrear as posições das pessoas ao longo do tempo, e explora-se a relação entre essas pessoas e suas orientações para detecção e classificação de grupos com base em informações sociológicas (*proxemics*, distâncias interpessoais, etc.). A geometria do grupo formado, assim como sua evolução ao longo do tempo, também são analisadas. Para a avaliação da atividade intragrupo, a evolução das áreas dos *blobs* correspondentes a cada pessoa é avaliada, para detecção de movimentos de gesticulação em grupos estáticos.

Como aplicações deste trabalho, pode-se mencionar a detecção e classificação automática de pequenos grupos em um *shopping center*, para que seja possível extrair padrões de comportamento segundo o estudo realizado.

Palavras-chave: Detecção de Grupo, Classificação de Grupo, Reconhecimento de Padrões, Acompanhamento de Objetos.

Abstract

The main goal of this work is to propose a model for an automatic (our semi-automatic) classification of groups using geometrical properties and for detecting intragroup activity based on video sequences. For group classification, a tracking algorithm is applied to obtain the position of each person across time, and the relationships among these people and their orientation are used to detect and classify groups based on sociological information (proxemics, interpersonal distance, etc). The geometry of the group, as well as its temporal evolution, are used to provide additional information on the group. To detect intragroup activity, the temporal evolution of blob areas related to tracked people is explored.

Regarding possible applications of this work, an example could be the detection and automatic classification of small groups in a shopping center, in order to extract behavior pattern according to this study.

Keywords: Group detection, Group classification, Pattern recognition, Tracking.

Lista de Figuras

1	Tramas existentes em pequenos grupos segundo (KEPHART, 1950)	15
2	Pares primitivos detectados no sistema segundo (HOSIE; VENKATESH; WEST, 1998)	17
3	Pessoas detectadas na cena através da identificação de movimentos segundo (HOSIE; VENKATESH; WEST, 1998)	18
4	Processo de acompanhamento e identificação de grupos no sistema segundo (MCKENNA et al., 2000)	19
5	Acompanhamento de pessoas através de matrizes combinadas com seus respectivos <i>blobs</i> segundo (FUENTES; VELASTIN, 2005)	20
6	Exemplo de aplicação do <i>tracker</i> definido no modelo de (DU et al., 2006) . .	20
7	Ilustração das distâncias interpessoais existentes segundo a classificação de (HALL, 1989)	23
8	Pares de robôs móveis em um esquema de controle distribuído segundo (YAMAGUCHI; ARAI; BENI, 2001)	26
9	Grupos de robôs móveis descrito por (YAMAGUCHI; ARAI; BENI, 2001) : a) grupo fechado - b) grupo aberto	26
10	Formação padrão linha segundo (IKEMOTO et al., 2005)	27
11	Formação padrão círculo segundo (IKEMOTO et al., 2005)	28
12	Formação padrão polígono segundo (IKEMOTO et al., 2005)	28
13	Simulação e demonstração da formação em colunas utilizando três robôs segundo (KRISHNANAND; GHOSE, 2005)	30
14	Diagramas de voronoi e espaço pessoal percebido (PPS)	35
15	Detecção e classificação de grupos segundo (JACQUES et al., 2007)	36
16	Extração da estrutura do grupo segundo características pré-definidas . . .	39
17	Extração de caraterísticas segundo a estrutura do grupo	39
18	Estruturação dos principais arquivos utilizados na modelagem do sistema .	44
19	Exemplo de aplicação do arquivo <i>automatic.m</i>	46
20	Imagem extraída na execução do arquivo <i>visu-voronoi-track.m</i>	46

21	Vetor - Identificação do Grupo - grupos formados por duas e três pessoas .	47
22	Seqüência de agrupamento formado por duas e três pessoas durante amostragem de vídeo	50
23	Visualizando os resultados graficamente	51
24	Detecção da interação entre as pessoas do grupo	51
25	Tela principal da interface web	52
26	Demonstração dos resultados extraídos do protótipo em Matlab através da interface web	56
27	Caso A: formação de um grupo composto por três pessoas	58
28	Formação de grupo envolvendo duas e três pessoas em distância íntima . .	60
29	Gráfico mostrando a variação da área dos <i>blobs</i> no decorrer do tempo para três pessoas	62
30	Identificação dos grupos detectados na cena no estudo de caso B	64
31	Identificação dos grupos no estudo de caso C	68
32	Ambiente de realização dos vídeos	72
33	Isolando variáveis no modelo de atenção	72
34	Comparação entre as variáveis distância e número de pessoas	73
35	Imagem da execução do modelo e extração de informações do grupo	74
36	Gráfico de distância do grupo em relação ao objeto de interesse ao longo do tempo.	75

Lista de Abreviaturas

DVD – Diagrama Dinâmico de Voronoi (*Dynamical Voronoi Diagram*)

PHP – Pré-Processador de Hipertexto (*Hypertext Preprocessor*)

PPS – Espaço Pessoal Percebido (*Perceived Personal Space*)

PS – Espaço Pessoal (*Personal Space*)

VD – Diagrama de Voronoi (*Voronoi Diagram*)

Sumário

1	Introdução	10
1.1	O Problema	11
1.2	Objetivos	12
2	Revisão Bibliográfica	14
2.1	Sistemas computacionais de detecção e acompanhamento de objetos, pessoas e grupos	15
2.2	Análise do comportamento humano baseado nas interações interpessoais . .	21
2.3	Análise de aplicações envolvendo grupos na Robótica	24
3	O Modelo Proposto	32
3.1	Fase 1: Detecção e Classificação do Grupo	33
3.2	Fase 2: Estrutura do Grupo	38
3.3	Fase 3: Atividade do Grupo	40
4	Protótipo	44
4.1	Mostrando resultados graficamente	49
4.2	Calculando as interações do grupo	50
4.3	Interface WEB para classificação das informações	51
5	Resultados Obtidos	57
5.1	Caso A: Encontro de três amigos	58
5.2	Caso B: Formação dinâmica de dois grupos	63
5.3	Caso C: Grupos compostos de duas e três pessoas em situações distintas .	67
6	Resultados Complementares	71
7	Conclusão	76

Capítulo 1

Introdução

Segundo (MILLS, 1970), o tema **grupo** já foi abordado em quase todos os ramos das ciências sociais. A relevância e o interesse no estudo dos pequenos grupos são facilmente compreendidos ao considerar-se que cada indivíduo integra, em média, cinco ou seis grupos em seu dia-a-dia, o que torna notável o aumento em pesquisa e reflexão a respeito de agrupamento.

O estudo sistemático dos pequenos grupos foi iniciado na década de 1950 (KEPHART, 1950). Desde então, muito foi feito, produzido e investigado em diversas áreas do conhecimento humano. Mas afinal, o que são os pequenos grupos? Para sintetizar de maneira simples, são *unidades compostas de duas ou mais pessoas que entram em contato, com um determinado objetivo e que consideram significativo este contato* (KEPHART, 1950).

Admite-se que o número de pequenos grupos seja surpreendente, mas porque estudá-los? Neste trabalho, foram escolhidos os pequenos grupos por duas razões principais: a razão sociológica e a razão computacional. A razão sociológica traz dados consistentes, que podem ser usados como base científica para o desenvolvimento de um modelo computacional. A tarefa é compreender pequenos grupos, fundamentados em conhecimento teórico, a respeito desses muitos bilhões de sistemas transitórios existentes no mundo.

A razão computacional é a que rege a aplicabilidade deste trabalho. Primeiramente, para que seja possível detectar um grupo em uma seqüência de vídeo e posteriormente classificá-lo, deve-se saber como o grupo se comporta, como é formado, qual a relação temporal existente na dinâmica do grupo, etc. Portanto, é necessário entender a formação do grupo, os tipos de distâncias existentes entre os indivíduos de um grupo, bem como as formações geométricas existentes em pequenos grupos e demais questões que envolvem a interação e atividade dos indivíduos que compõem o grupo.

Os pequenos grupos constituem um caso especial do tipo mais geral de sistema social. Não são apenas microssistemas, mas são também, fundamentalmente, microcosmos de sociedades mais amplas. Através do exame cuidadoso desses microssistemas é possível construir modelos computacionais, que aplicados às seqüências de vídeos existentes, permitam verificar a veracidade dos pequenos grupos, tais como: sistemas de detecção de atividade suspeita, monitoramento de grupos, vigilância, entre outros.

1.1 O Problema

Tendo em vista os vários aspectos que envolvem o estudo de grupos, focalizou-se este trabalho na classificação automática de grupos quanto às interações entre os membros que compõem o grupo e também nas atividades individuais dos mesmos. O estudo de grupos em processos automáticos computacionais é um trabalho desafiador, devido a pouca bibliografia sobre detecção/classificação de grupos, sendo necessário realizar um estudo exploratório sobre a relação entre a natureza do grupo formado e a distribuição espacial dos elementos do grupo.

Uma maneira de extrair as posições das pessoas automaticamente para detecção e classificação de grupos é conseguida através do algoritmo de *tracking* proposto por (JACQUES, 2006) o qual faz uso de técnicas de visão computacional. Neste aspecto, deve-se levar em consideração que o resultado dos algoritmos de acompanhamento baseados em

visão computacional pode gerar trajetórias erradas ou deixar de acompanhar um determinado indivíduo, sendo necessária uma análise detalhada com relação ao acompanhamento dos indivíduos que compõem o grupo no decorrer da cena.

A classificação de atividades de cada membro do grupo e/ou do grupo como um todo, juntamente com a identificação da geometria formada pelo grupo que está sendo estudado, são os grandes desafios desta pesquisa, o que leva a fazer um estudo científico exploratório desde a parte referente a formação de um grupo (sociologia), seus aspectos comportamentais (distâncias interpessoais) e seus aspectos visuais (geometria do grupo) contribuindo para um trabalho completo e inovador.

1.2 Objetivos

O objetivo principal deste trabalho é desenvolver um modelo que permita a classificação automática (para grupos em movimento) ou semi-automática (para grupos em estado estacionário) de grupos de pessoas, quanto aos seus níveis de interação e atividade, utilizando informações de geometria e detecção de atividades intragrupo, a partir de uma seqüência de vídeo.

Este trabalho foi desenvolvido seguindo dois aspectos fundamentais: a detecção da atividade individual de cada componente do grupo e a classificação dos grupos conforme definido a seguir:

Detecção da interação e atividade de grupo: detecção nas seqüências de vídeo da atividade individual de cada membro que compõe o grupo. O isolamento desta informação é de suma importância para o trabalho como um todo, pois uma possível aplicação desta classificação é a identificação de qual (ou quais) indivíduo(s) está(ão) mais ativo(s) em um grupo, informação que pode ser repassada para um sistema de registro automático de fotografias, por exemplo.

Classificação de grupos: estudo exploratório sobre agrupamentos, ou seja, a

detecção e classificação do grupo, a quantidade de pessoas envolvidas num grupo e também a geometria que o grupo forma. Sabe-se que a geometria de um grupo está diretamente associada ao número de componentes que compõem o mesmo (KEPHART, 1950), sendo que grupos maiores tendem a ter um número maior de tramas entre os membros, crescendo gradativamente conforme aumenta o número de componentes.

A organização deste trabalho está distribuída conforme a estrutura a seguir: no capítulo 2, tem-se a revisão bibliográfica com relação ao tema proposto neste trabalho; o capítulo 3, apresenta a estrutura do modelo proposto a ser desenvolvido; o capítulo 4, mostra de maneira geral o protótipo desenvolvido; no capítulo 5, são mostrados os resultados obtidos no trabalho, envolvendo grupos de pessoas em situações diversas; no capítulo 6, tem-se os resultados complementares, obtidos através da análise de dados envolvendo um modelo de atenção de grupos. Por fim, o capítulo 7 contém as conclusões finais.

Capítulo 2

Revisão Bibliográfica

Nesta revisão bibliográfica, procurou-se buscar diversos trabalhos no meio científico que fossem relacionados com o tema abordado neste trabalho. A seguir, são apresentados vários trabalhos que estão sendo desenvolvidos, os quais servem de base para este projeto, tais como: detecção de indivíduos, grupos e multidões; acompanhamento de entidades móveis, artigos relacionados ao estudo de grupos, sua formação propriamente dita, onde se buscou informações relevantes na sociologia (ciências sociais), além de comportamentos de grupos aplicados a robótica.

Neste estudo exploratório começa-se com a essência do grupo apresentado por (KEPHART, 1950) que descreve informações sobre grupos de pessoas e suas relações. O autor relata que num grupo de pessoas, o tamanho ou número de tramas que vem a se desenvolver pode ser determinado tanto pelo número de membros que o compõe quanto por uma relação mais completa onde são levados também em conta o número e a posição dos vínculos de relação interindivíduo. Se existem vínculos entre todos os membros de um grupo de n membros, haverá $\frac{1}{2}n(n-1)$ vínculos. Ao considerar o resultado de um aumento no número de membros num grupo, em que todos os membros podem falar livremente, o número dos canais de comunicação que será necessário manter pode ser calculado por essa mesma fórmula. Um fator que contribui para uma grande complexidade dos relacionamentos nos grupos é o fato de que o número de vínculos aumenta naturalmente com o

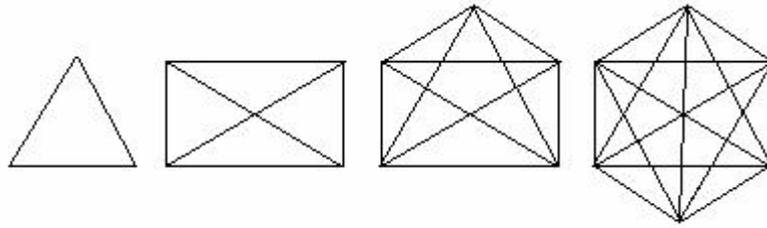


Figura 1: Tramas existentes em pequenos grupos segundo (KEPHART, 1950)

aumento do número de membros. O acréscimo de um membro exige tantos novos vínculos quantos eram os membros antes desse acréscimo.

Por exemplo, num grupo de 3 membros, o número de vínculos é 3. Num grupo de 4 membros, o número de vínculos é 6. Num grupo de 5 membros, o número de vínculos é 10, e assim sucessivamente, conforme ilustrado na Figura 1.

A seguir apresenta-se alguns trabalhos relatados na literatura, que descrevem em sua ordem: sistemas computacionais de detecção e acompanhamento de objetos, pessoas e grupos, conforme a seção 2.1, que servem como base para o trabalho desenvolvido, devido ao fato deste trabalho utilizar técnicas extraídas de visão computacional; alguns trabalhos sobre a análise do comportamento humano baseado nas interações interpessoais, apresentados na seção 2.2, e por fim, na seção 2.3, os trabalhos com relação a análise de aplicações envolvendo grupos na robótica, onde o objetivo é relacionar as técnicas utilizadas para simular grupos com robôs e aplicá-las em situações reais, onde existam formações de grupos de pessoas.

2.1 Sistemas computacionais de detecção e acompanhamento de objetos, pessoas e grupos

Existem inúmeros trabalhos desenvolvidos na área de detecção e acompanhamento de objetos (YILMAZ; JAVED; SHAH, 2006). Estes visam detectar diversas entidades, dentre elas, pessoas, grupos, multidões, faces, partes do corpo, objetos, entre outros. Como a

tecnologia de acompanhamento não é o escopo de contribuição desta pesquisa, somente alguns poucos trabalhos serão apresentados.

Segundo (WANG; HU; TAN, 2003) a análise visual do movimento humano busca detectar, acompanhar e entender o comportamento das pessoas em seqüências de vídeos. Tais análises têm despertado grande interesse em pesquisas envolvendo visão computacional, devido a abrangência de suas aplicações, como por exemplo, vigilância, extração de padrões de comportamentos, entre outros.

A detecção tem por objetivo segmentar as regiões correspondentes às pessoas em relação ao restante da imagem. Este é um passo fundamental para sistemas de análise do movimento humano, já que as etapas subseqüentes como acompanhamento e entendimento do comportamento humano são bastante dependentes da detecção.

O processo de detecção normalmente envolve segmentação dos objetos em movimento e sua classificação. Por exemplo, imagens capturadas por uma câmera de vigilância posicionada dentro de uma universidade podem incluir pessoas, pássaros, nuvens e outros objetos em movimento. Para futuramente acompanhar as pessoas, é necessário distingüí-las corretamente dos outros objetos. Assim sendo, essa etapa de classificação, pode não ser necessária em algumas situações, por exemplo, quando se sabe que os objetos em movimento tratam-se somente de pessoas (WANG; HU; TAN, 2003). O acompanhamento de indivíduos em seqüências de vídeo é um passo fundamental em visão computacional, para se entender a dinâmica do comportamento humano. O objetivo principal é acompanhar o movimento das pessoas em uma seqüência de quadros do vídeo.

O resultado do acompanhamento pode ser analisado para se interpretar o comportamento do indivíduo em estudo (atividade, interação, geometria formada, variação das posições no decorrer do tempo, etc.). O acompanhamento no decorrer do tempo envolve normalmente estabelecer uma relação coerente para um objeto em quadros consecutivos do vídeo, usando características como pontos, linhas ou regiões.

Com relação aos sistemas de detecção e classificação de comportamento de gru-

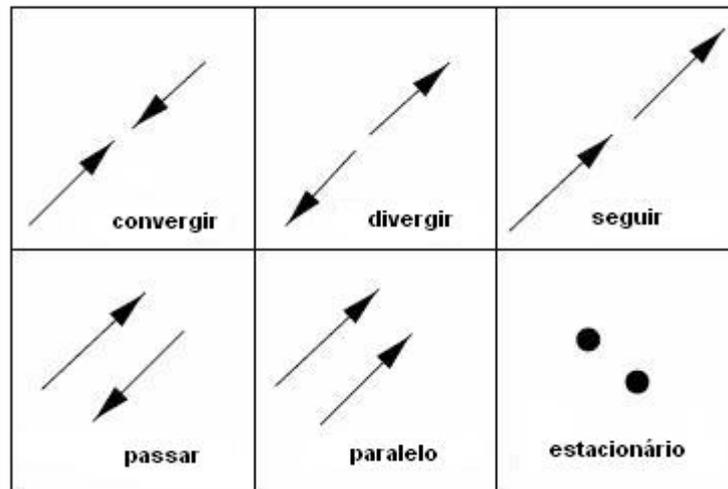


Figura 2: Pares primitivos detectados no sistema segundo (HOSIE; VENKATESH; WEST, 1998)

pos tem-se o trabalho de (HOSIE; VENKATESH; WEST, 1998) o qual aborda a detecção e classificação do comportamento de grupos sobre dados de inspeção visual. Testes padrão simples são usados como blocos para descrever o comportamento sobre uma seqüência temporal, ou comparados com pares primitivos de movimentos, para detectar o comportamento do grupo. Pares primitivos formam um conjunto definido de movimentos que poderiam ocorrer entre dois alvos na cena sobre uma amostragem de tempo. Neste sistema, os pares primitivos treinados para detecção são mostrados na Figura 2.

Estas disposições de pares foram selecionadas com o intuito de descrever uma variedade larga de movimentos do grupo em uma situação humana de vigilância. O sistema faz uso de árvore de decisão para distinguir entre cada classe de pares primitivos. Uma consideração importante é o número de pessoas que se quer identificar na cena. Se a tarefa for detectar somente as ações exibidas por indivíduos, então a solução é relativamente simples, pois basta analisar cada movimento separadamente. Entretanto, no exemplo de detecção do comportamento de grupo, a tarefa requer segmentação dos dados relativos as pessoas, a fim de encontrar significativos relacionamentos entre elas. O sistema tem por objetivo principal detectar de maneira robusta grupos de pessoas que possuem movimentos similares, conforme pode ser visto na Figura 3.

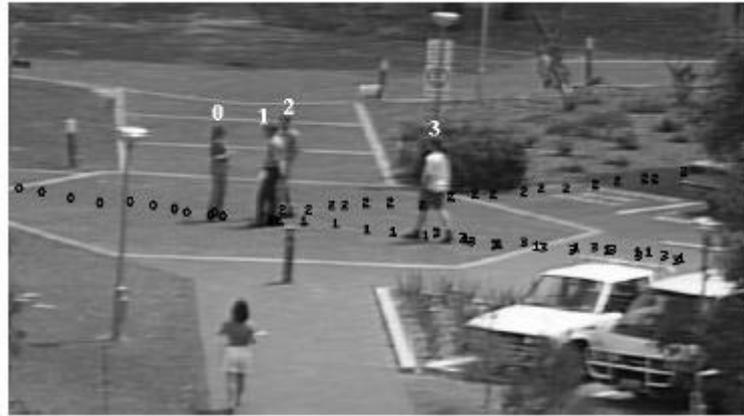


Figura 3: Pessoas detectadas na cena através da identificação de movimentos segundo (HOSIE; VENKATESH; WEST, 1998)

O trabalho descrito por (MCKENNA et al., 2000) trata de acompanhamento aplicado a grupos de pessoas. O autor descreve um sistema de visão computacional para realizar acompanhamento de múltiplas pessoas em ambientes livres. O acompanhamento é executado em três níveis de abstração: regiões, pessoas e grupos.

- Regiões: são os componentes conexos que foram seguidos por uma quantia mínima de quadros.
- Pessoas: uma pessoa consiste de uma ou mais regiões agrupadas. Para cada pessoa é definido um modelo de aparência.
- Grupos: um grupo consiste de uma ou mais pessoas agrupadas. Se duas pessoas formam uma região, elas são consideradas como estando no mesmo grupo.

As pessoas da cena são seguidas com suas oclusões mútuas enquanto dão formas a grupos. As interações simples com objetos podem também ser detectadas no sistema proposto e posteriormente testadas em seqüências de vídeo em ambientes fechados e abertos.

A estruturação do sistema procede da seguinte maneira (conforme pode ser visto na Figura 4): a primeira imagem (a) mostra o resultado da subtração do fundo. A imagem ao lado (b) mostra o resultado da subtração do fundo usando gradientes; a seguir é mostrado o resultado da subtração combinada do fundo (c); na imagem (d) é mostrado o

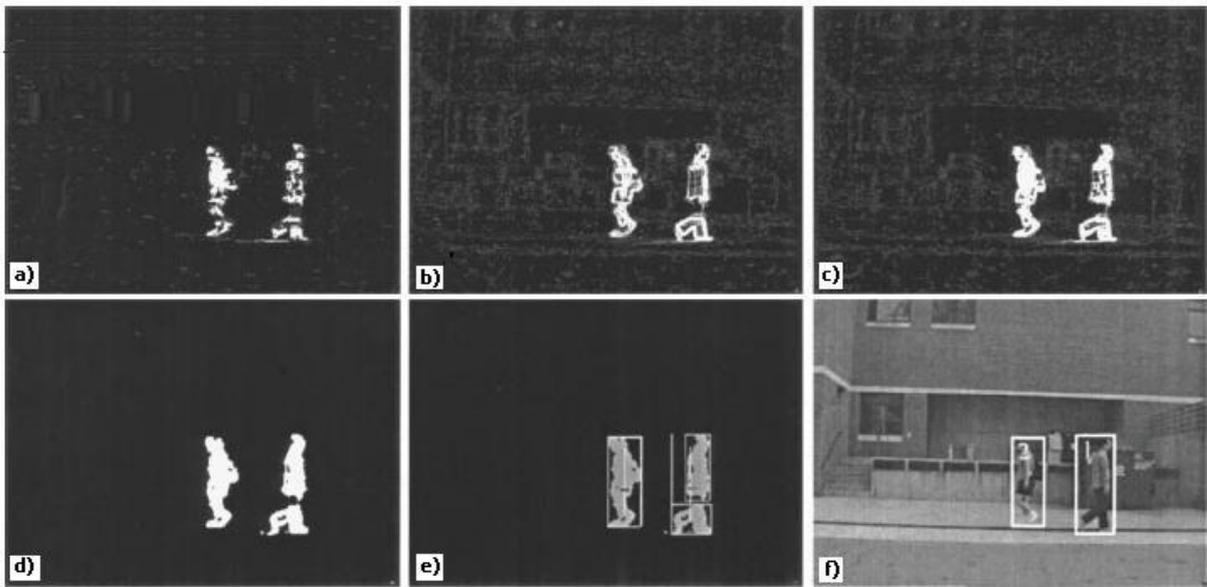


Figura 4: Processo de acompanhamento e identificação de grupos no sistema segundo (MCKENNA et al., 2000)

resultado da aplicação do filtro mediano; em seguida, imagem (e) são aplicadas as caixas limitadoras para as regiões e pessoas que estão sendo seguidas na cena, (uma pessoa consiste de duas regiões) e por fim, imagem (f) as caixas limitadoras aplicadas na imagem real e acompanhamento das pessoas detectadas na cena.

No trabalho de (FUENTES; VELASTIN, 2005), o método desenvolvido foi o acompanhamento de pessoas para aplicações de vigilância. O autor apresenta um algoritmo que permite seguir múltiplos objetos em ambientes complexos. Os *pixels* do primeiro plano são detectados usando o contraste de luminosidade e agrupamento das pessoas em *blobs*, que são acompanhados no tempo.

O autor propõe também algumas aplicações em sistemas automáticos de vigilância, ligando informações relativas a posição das trajetórias e do *blob*, com os eventos a serem detectados em uma seqüência de vídeo. As posições das pessoas podem ser usadas para detectar eventos baseados em posições, tais como intrusão em áreas proibidas, etc. Um exemplo da aplicação do algoritmo proposto pode ser visto conforme mostra a Figura 5.

No trabalho desenvolvido por (DU et al., 2006) foi apresentado um modelo para identificar interações entre pessoas. Para tal, as características de cada pessoa são de-



Figura 5: Acompanhamento de pessoas através de matrizes combinadas com seus respectivos *blobs* segundo (FUENTES; VELASTIN, 2005)

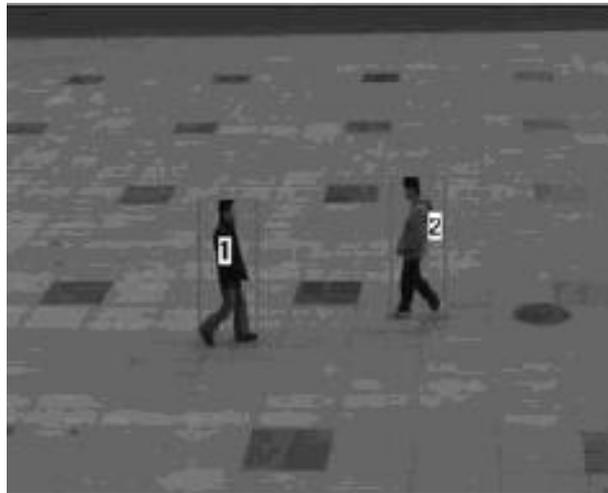


Figura 6: Exemplo de aplicação do *tracker* definido no modelo de (DU et al., 2006)

terminadas em dois conjuntos: características globais e locais. As globais descrevem as características do próprio movimento da pessoa e as características do movimento entre pessoas, como velocidade, direção, distância e ângulo entre duas pessoas. As locais descrevem características individuais como contorno, movimento, ângulo do tronco, largura, comprimento do objeto e outros.

A modelagem das atividades de interação é feita através de Redes de Bayes. O modelo utiliza um tempo de treinamento para alimentar a Rede de Bayes, a qual é criada utilizando-se de parâmetros estimados inicialmente.

No modelo de (DU et al., 2006) os comportamentos procurados são fixos, ou seja, eles

são reproduzidos para treinamento da Rede de Bayes e depois os mesmos são avaliados pela rede. Portanto, o usuário não pode especificar um novo comportamento para ser procurado posteriormente.

No trabalho de (JACQUES, 2006) foi desenvolvido um algoritmo para acompanhamento de pessoas, cujo principal objetivo foi utilizar-se de técnicas de visão computacional para simular comportamentos de multidões de humanos virtuais.

Resumidamente, o algoritmo recebe um vídeo de entrada, em escalas de cinza (a utilização de seqüências de vídeo monocromáticas confere uma maior flexibilidade ao modelo) e remove o *background* da cena, deixando apenas os *pixels* do *foreground* (*blobs*). Como o algoritmo de (JACQUES, 2006) trabalha com câmeras de topo, a perspectiva das cabeças das pessoas detectadas na cena é praticamente invariante, e um algoritmo de correlação de máscaras é aplicado para acompanhar uma pessoa em quadros consecutivos. Assim, para cada pessoa i , a sua trajetória é formada por uma lista de pontos $(x_i(t), y_i(t))$, que denota sua evolução ao longo do tempo (ou seja, a trajetória da pessoa).

Salienta-se ainda que existem diversas outras abordagens para o problema de acompanhamento de objetos, como (BELEZNAI; FRÜHSTÜCK; BISCHOF, 2004), (COLOMBARI; FUSIELLO; MURINO, 2007), (CUI et al., 2007), (KANG; KIM; BANG, 2002), (MARCENARO; MARCHESOTTI; REGAZZONI, 2006), (WANG; D., 2007). Entretanto, este trabalho adotou o método desenvolvido por (JACQUES, 2006), pois foi implementado para trabalhar com imagens de topo, as quais seguem o mesmo padrão neste trabalho.

2.2 Análise do comportamento humano baseado nas interações interpessoais

As distâncias interpessoais são medidas com grande potencial para caracterização de grupos. Entendê-las e interpretá-las é um passo crucial no estudo do comportamento de grupos (SOMMER, 1973). A distância mantida entre duas pessoas, durante um diálogo,

faz parte da linguagem corporal e pode mudar conforme a situação vivenciada.

Segundo (HALL, 1989), que estuda há vários anos o comportamento humano, as distâncias interpessoais podem ser divididas em quatro categorias principais: distância íntima, distância pessoal, distância social e distância pública.

- **Distância Íntima**

- *Distância íntima próxima*: contato - esta distância é muito natural entre pessoas com relações íntimas, em que há uma grande consciência da proximidade. É natural entre crianças e pais, casais, namorados, etc.
- *Distância íntima afastada*: de 15 a 50 cm - é a distância aceitável para cumprimentos formais (apertos de mão, entre outros).

- **Distância Pessoal**

- *Distância pessoal próxima*: de 50 a 80 cm - é a distância normal para cumprimentos formais. É uma distância confortável em reuniões sociais.
- *Distância pessoal afastada*: de 80 a 120 cm - distância normal para uma discussão pessoal, pois confere privacidade à comunicação. Quando a conversa convida à intimidade, a distância pode ser reduzida, mas tendo o cuidado de utilizar uma linguagem adequada a esse grau de intimidade.

- **Distância Social**

- *Distância social próxima*: de 1,2 a 2 metros - distância adequada a negócios interpessoais. É a distância normal entre cliente e vendedor e patrão e empregado, pois evita familiaridades.
- *Distância social afastada*: de 2 a 3 metros - é a distância considerada normal para relações formais. É a esta distância a que podemos colocar-nos de uma

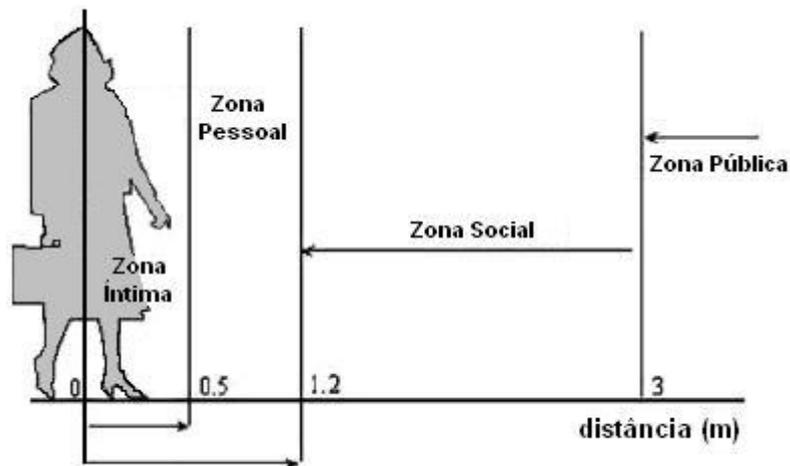


Figura 7: Ilustração das distâncias interpessoais existentes segundo a classificação de (HALL, 1989)

pessoa que esteja a trabalhar, sem a obrigarmos a interromper o seu trabalho para comunicar-se conosco.

• Distância Pública

- *Distância pública próxima:* de 3 a 8 metros - distância natural em reuniões informais, aulas, entre outros.
- *Distância pública afastada:* mais do que 8 metros - distância adequada para aulas e grandes reuniões. É a distância de excelência para os políticos e para os atores, pois os gestos podem ser mais estilizados e mais simbólicos.

Sabendo das distâncias interpessoais existentes, pode-se visualizá-las conforme mostra a Figura 7.

Segundo (SOMMER, 1973) existe uma considerável semelhança entre *espaço pessoal* e *distância individual*, ou espaçamento característico dos membros da espécie. A distância individual existe apenas quando dois ou mais membros da espécie estão presentes e sofre grande influência da densidade da população e do comportamento territorial. A distância pessoal e o espaço pessoal interagem em sua influência sobre a distribuição de pessoas. A

violação da distância individual é uma violação das expectativas da sociedade, a invasão do espaço pessoal é uma intrusão nas fronteiras do “eu” da pessoa.

O espaço pessoal está relacionado com o nível de conforto das pessoas. Mesmo que seja intuitivo pensar que uma pessoa sinta-se confortável se o espaço ao seu redor é suficientemente grande, acredita-se que o espaço pessoal não seja a melhor métrica para estimar o nível de conforto (JACQUES et al., 2007). De fato, o conceito psicológico de espaço pessoal é estático, significando que é considerada a região ao redor do indivíduo, desconsiderando sua direção e sentido de movimento. Uma pessoa pode ter um espaço consideravelmente grande atrás de si, porém pode não sentir-se confortável quando existe outra pessoa ou um obstáculo a sua frente (HALL, 1989).

2.3 Análise de aplicações envolvendo grupos na Robótica

No trabalho de (MATARIC, 1995), foram definidos e caracterizados os princípios básicos no projeto de agentes autônomos organizados em grupos, definindo conceitos (relativo às origens dos grupos na robótica), juntamente com os principais comportamentos de grupos e suas respectivas aplicações.

O problema de sintetizar e analisar agentes autônomos em grupos tem sido estudado a pouco mais de 10 anos, na área da robótica. Por isso, faz-se necessário, caracterizar os comportamentos existentes em grupos de robôs, bem como definir a origem destes grupos.

Comportamento e Objetivos

No que diz respeito a comportamentos e seus objetivos, define-se comportamento como sendo uma lei de controle para alcançar e/ou manter um objetivo particular (MATARIC, 1995). Por exemplo, no domínio de um robô, “seguir” é uma lei de controle

que faz exame das entradas dos sensores de um agente e faz gerar as ações que manterão o agente se movendo dentro de uma região fixa, atrás de um outro objeto móvel.

Ao invés de tentar analisar comportamentos complexos arbitrários, focalizou-se em fornecer um conjunto de comportamentos primitivos que podem ser usados para sintetizar e analisar um tipo particular de sistemas complexos multiagentes. Os grupos primitivos fornecem uma linguagem para projetar análise no controle dos programas e resultam em comportamentos de grupos, tais como: agrupamento, procura por outros robôs, etc.

As teorias básicas do comportamento são geralmente extraídas da biologia, tais como: seguir algo, dispersar, agrupar, etc (MATARIC, 1995). Estas teorias são construídas segundo comportamentos observados em uma variedade de espécies na natureza. Conseqüentemente, pode ser aplicável em vários domínios da interação presentes em sistemas multiagentes. Esta metodologia é constante, independentemente do tamanho do grupo e dos tipos de interações.

No trabalho de (YAMAGUCHI; ARAI; BENI, 2001) foi descrito um sistema de controle distribuído, para realizar formações de grupos com robôs. Uma das funções mais importantes em que os sistemas são requeridos, quanto à execução de tarefas, é a potencialidade em coordenar comportamentos dos multiagentes, para que eles trabalhem de forma cooperativa. Os robôs multiagentes em cooperação são distribuídos em espaços bidimensionais e tridimensionais (em alguns casos), e trabalham juntos, seguindo alguns comandos dados por um supervisor, em um esquema centralizado de controle ou seguindo regras projetadas para aplicações avançadas, em um esquema distribuído de controle.

O objetivo principal do trabalho foi mostrar que existe um esquema distribuído de controle de multiagentes, em que:

- o movimento de cada robô é determinado por fórmulas matemáticas, incluindo o vetor da formação de grupos de forma quantitativa;
- todos os robôs são estabilizados, quanto ao movimento;

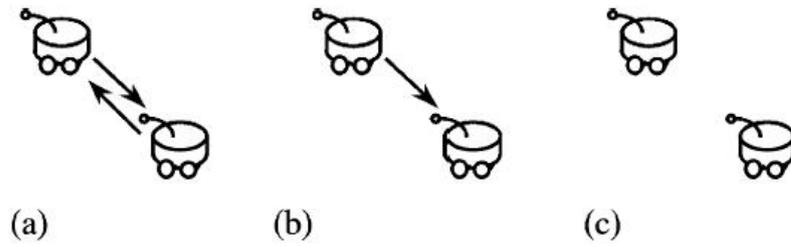


Figura 8: Pares de robôs móveis em um esquema de controle distribuído segundo (YAMAGUCHI; ARAI; BENI, 2001)

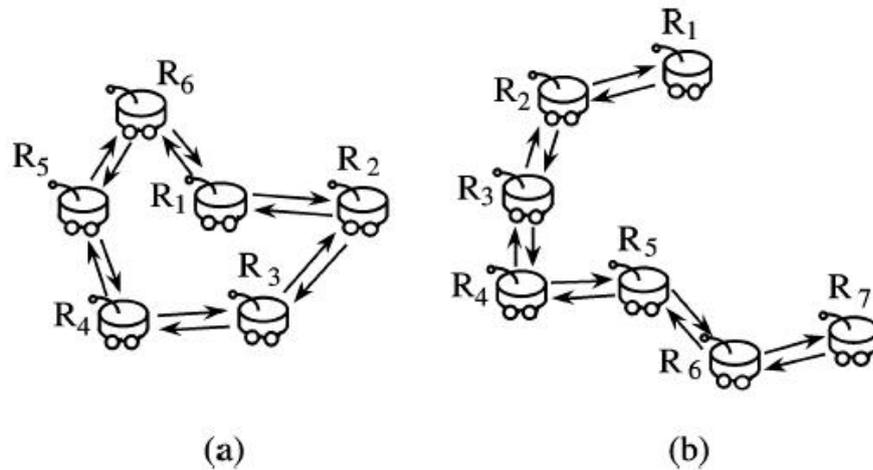


Figura 9: Grupos de robôs móveis descrito por (YAMAGUCHI; ARAI; BENI, 2001) : a) grupo fechado - b) grupo aberto

- a formação é controlada pelos vetores da mesma.

Foram descritos os sentidos de cada robô, sua posição relativa e sua orientação relativa a outro robô, de forma que qualquer par de robôs esteja em uma das três relações: (a) ambos detectam, (b) somente um detecta, (c) nenhum robô detecta outro, conforme mostra a Figura 8. As setas presentes na Figura 8 vão do robô que está detectando para o robô que é detectado. Foram classificados os grupos de robôs móveis em dois tipos de configuração (*a* e *b*), conforme a Figura 9. O grupo (a), tem uma configuração denotada como “um grupo fechado” enquanto que o grupo (b) tem uma configuração denotada como “um grupo aberto”.

No experimento proposto no trabalho, dois robôs tentam manter distâncias relativas diferentes, se movendo em um plano com suas posições controladas e tendo um



Figura 10: Formação padrão linha segundo (IKEMOTO et al., 2005)

controle da sua formação de grupo com relação a outros robôs. O sistema distribuído garante a estabilidade e o controle dos agentes, pela informação local que possui de cada um dos agentes. Cada agente tem seu próprio sistema de coordenadas, detectando sua posição e orientação com relação a outro agente. A validação deste sistema é suportada através de simulações de computador.

No trabalho de (IKEMOTO et al., 2005) foi proposto um algoritmo para formação gradual de padrões, em sistemas de múltiplos robôs homogêneos. O algoritmo permite que um grupo de robôs, gere de forma gradual, alguns padrões de formações de grupos complicados, tais como um polígono, através de padrões simples (linhas, colunas, etc.), usando somente informações locais que um robô pode obter. As formações padrões testadas nos robôs são: linhas, círculos e polígonos.

Formação padrão linha

Para realizar a formação linha, múltiplos robôs ficam emitindo infra-vermelho em seus movimentos traseiros, de forma aleatória em um campo. Se um robô detecta o infra-vermelho de outro robô, então ele começa a segui-lo. Se outro robô se junta a linha, este envia uma mensagem ao robô que está a sua frente. O robô do alto da linha pode saber o comprimento da linha contando as mensagens enviadas pelos robôs subsequentes. Esta representação pode ser vista conforme Figura 10.

Formação padrão círculo

O padrão círculo é formado pelo robô de topo de uma linha, movimentando-se em um formato de arco. Os robôs situados na linha seguem apenas o robô do topo,



Figura 11: Formação padrão círculo segundo (IKEMOTO et al., 2005)

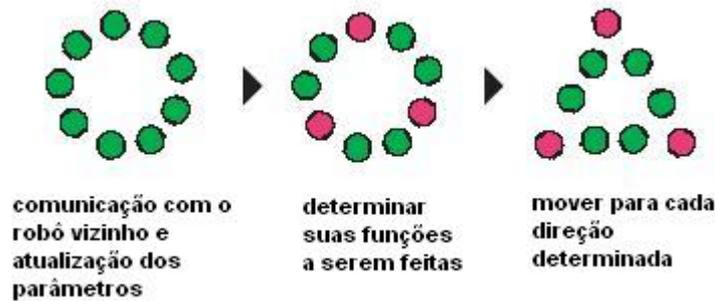


Figura 12: Formação padrão polígono segundo (IKEMOTO et al., 2005)

controlando suas velocidades e sentido a fim de receber uma intensidade constante de infra-vermelho. Embora o raio da curvatura gerado dependa do número de robôs, a curvatura é ajustada baseada no número de robôs contados na formação padrão linha, conforme pode ser visto na Figura 11.

Formação padrão polígono

Para formação de um polígono, cada robô comunica o valor de seus fatores com os dois robôs mais próximos (o robô da esquerda e o da direita). Quando um robô recebe o valor dos fatores de seus vizinhos, atualiza os seus valores. Após um determinado número de emissões, cada robô envia uma mensagem a todos os robôs para terminar a atualização dos valores. Assim se um robô tiver o valor maior que seus vizinhos, fica fazendo o papel de ângulo de um polígono, caso contrário, se os valores forem menores que os dos vizinhos, movem-se para dentro do círculo, fazendo o papel de borda, para formação do polígono, conforme pode ser visto na Figura 12.

Além das formações padrões simples, outras formações foram testadas tais como,

quadrado e hexágono, tornando possíveis as transformações de padrões simples em padrões mais complexos, através do algoritmo proposto por (IKEMOTO et al., 2005).

Segundo (KRISHNANAND; GHOSE, 2005) as formações de grupos de robôs têm sido estudadas recentemente, e podem possuir diversas aplicações, tais como: planejamento do movimento de robôs, desvio de obstáculos, estabilidade de robôs, métodos de formações de grupos, etc. Os métodos da formação de grupos, por sua vez, podem ser classificados em três categorias: líder-seguidor, estrutura virtual e bases de comportamento, sendo que neste trabalho o autor trabalha com bases de comportamento. As bases de comportamento, por sua vez, são esquemas descentralizados na natureza e atendem bem às exigências das formações de grupos com grande número de agentes.

Inicialmente, um espaço de trabalho bidimensional é criado e todos os robôs são distribuídos de forma aleatória ou colocados em posições definidas pelo usuário no espaço de trabalho. Os módulos de percepção, na seleção do comportamento e da atuação, são modelados algoritmicamente. Em cada iteração, os vários comportamentos do robô são selecionados baseados no estado detectado no módulo da percepção. Os robôs podem ser programados para comportar uma série de formações específicas, dentre delas a formação em colunas, conforme pode ser visto na Figura 13.

Existem ainda outros trabalhos envolvendo grupos na robótica, que servem de referência a este trabalho de pesquisa, dentre eles pode-se citar: (PAYTON; ESTKOWSKI; HOWARD, 2003), (MATARIC, 1994), (BARFOOT; CLARK, 2003) e (REIF; WANG, 1995).

Nos trabalhos citados com relação a análise de aplicações envolvendo grupos na robótica, o objetivo principal é mostrar os trabalhos que desenvolvem formações de robôs, utilizando-se de técnicas para classificação dos robôs segundo a geometria formada. As diferentes formações geométricas realizadas pelos robôs pode ser o ponto de partida para detecção de novos padrões, além de servirem como experiência inicial para simulação de grupos maiores a serem testados.

A utilização de robôs para simular humanos pode trazer algumas vantagens, nas

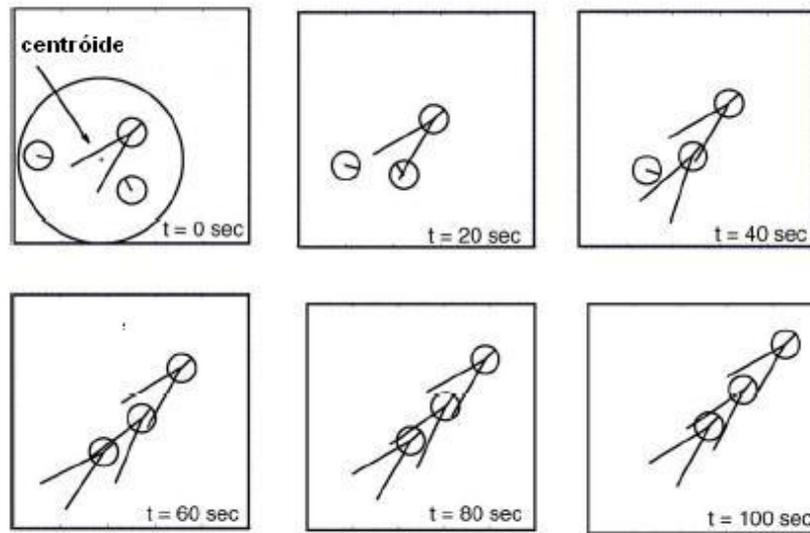


Figura 13: Simulação e demonstração da formação em colunas utilizando três robôs segundo (KRISHNANAND; GHOSE, 2005)

quais pode-se citar:

- simular diferentes tipos de grupos, alterando suas características, tais como: velocidade, direção, detecção de outro indivíduo, entre outros;
- realizar experimentos controlados e de objetivos específicos;
- ter um controle centralizado do experimento ou um controle distribuído das tarefas que serão realizados por um grupo de robôs;

Com relação as desvantagens em utilizar-se robôs para simular grupos de pessoas, tem-se:

- imprecisão dos movimentos humanos;
- custos de implementação (aumentando gradativamente dependendo do número de robôs a ser utilizado no experimento);

Algo bastante interessante tanto para este trabalho como para trabalhos futuros é com relação as teorias básicas do comportamento, geralmente extraídas da biologia e que podem ser aplicadas em conjunto tanto para testes com robôs como para humanos,

tais como: seguir algo, dispersar, agrupar, conforme citado no trabalho de (MATARIC, 1995). Tais teorias do comportamento, podem ser estudadas com o objetivo de formar novos padrões de comportamento, que venham a contribuir para resolver um determinado problema, por exemplo.

Os trabalhos utilizando robôs bem como os trabalhos que utilizam-se de humanos, complementam-se na medida em que existem aplicações que são mais vantajosas utilizando-se de robôs, como é o caso de aplicações de alto risco, bem como situações onde existe perigo de vida. Já outras aplicações como o teste de um padrão detectado na robótica deve ser utilizado com humanos para saber se realmente fazem sentido as informações detectadas pelos testes com robôs.

A evolução de trabalhos voltados para a detecção e classificação de padrões de grupos utilizando robôs pode ser associado, a trabalhos que utilizam-se de visão computacional, para prover de forma automática a classificação e detecção de atividades dos indivíduos que compõem um grupo, como é o caso de pesquisa deste trabalho.

O modelo proposto desenvolvido neste trabalho é apresentado a seguir.

Capítulo 3

O Modelo Proposto

Analisados os principais trabalhos nesta área (capítulo 2), tem-se duas linhas de pesquisas distintas para o trabalho que está sendo desenvolvido. A primeira diz respeito a trabalhos desenvolvidos no escopo de visão computacional (detecção, acompanhamento, reconhecimento de objetos, pessoas e grupos), e a segunda, referente ao estudo de grupos, aplicando modelos comportamentais relatados na literatura. Assim, como contribuição deste trabalho, busca-se desenvolver um modelo para detectar grupos de pessoas (com ou sem movimento) em seqüências de vídeo e classificá-los segundo algumas características, tais como: estrutura do grupo (polígono formado, área e variação da área), atividade do grupo (interações presentes e variação dos *blobs* individuais dos agentes) bem como mudanças nas estruturas do grupo.

A primeira tarefa a ser executada neste trabalho é compor um banco de dados de vídeos diversos, ou seja, ter em mãos vídeos com diferentes formações de pessoas (grupos de duas e três pessoas) dos quais possam ser extraídas as informações que são necessárias neste trabalho. Para isso, utilizam-se vídeos diversos, coletados e utilizados pelo laboratório CROMOS (<http://www.inf.unisinos.br/cromoslab>), da Unisinos/RS, para coleta dos dados propostos e realização de testes do modelo proposto.

A seguir apresentam-se as fases do modelo proposto, correspondente as etapas que

são realizadas para a extração das características dos grupos e aplicação do modelo: a **fase 1** é responsável pela detecção e classificação do grupo (grupos voluntários e involuntários); a **fase 2** é referente a estrutura do grupo (área formada, geometria e variação da área no tempo); a **fase 3** é relativa a atividade do grupo (interações presentes e variação dos *blobs* individuais); e, por fim, a **fase 4** é caracterizada quando ocorre um novo evento no sistema (mudança de estrutura do grupo), fase esta responsável pela reestruturação e reclassificação do grupo.

3.1 Fase 1: Detecção e Classificação do Grupo

A primeira fase do modelo proposto consiste em detectar grupos de pessoas em seqüências de vídeos e posteriormente fazer a sua classificação. A detecção de pessoas em seqüências de vídeo é feita através do algoritmo desenvolvido por (JACQUES, 2006) que visa separar o *background* (fundo) da cena dos objetos de *foreground* (frente), de modo a identificar o objeto que está em movimento. O algoritmo de (JACQUES, 2006) é baseado na correlação de máscaras usando a soma dos quadrados das diferenças como métrica de casamento. Posteriormente, a pessoa detectada é mapeada através de uma área delimitada denominada *blob* e acompanhada durante a cena (*tracking*).

Para que seja possível extrair informações referentes a classificação do grupo detectado, usa-se o modelo desenvolvido por (JACQUES et al., 2007), que provê uma classificação de pessoas e grupos seguindo os seguintes itens: pessoas confortáveis e desconfortáveis, grupos voluntários e involuntários. Os conceitos de diagrama de voronoi (VD), espaço pessoal (PS) e diagrama dinâmico de voronoi (DVD) são abordados para um melhor entendimento do processo de classificação de grupo.

A posição de cada pessoa detectada na cena é usada como um sítio para computar o diagrama de voronoi (VD) em cada quadro. A evolução temporal do VD leva ao conceito de diagrama dinâmico de voronoi (DVD), que é explorado visando acompanhar/armazenar

a variação dos parâmetros sociológicos e psicológicos (HALL, 1989) ao longo do tempo. Estes parâmetros são, basicamente, o espaço pessoal de cada indivíduo e a distância entre eles, os quais são usados para reconhecer eventos, formação de grupos e outras características (JACQUES et al., 2007). Vale lembrar que as posições de cada pessoa no tempo são obtidas através de um algoritmo de visão computacional em seqüências de vídeo reais, neste trabalho as posições de cada pessoa foram obtidas através do algoritmo desenvolvido por (JACQUES, 2006).

O conceito de espaço pessoal (PS) é definido como uma área invisível ao redor do corpo de um indivíduo. Esta área é relatada como uma zona de conforto durante a comunicação interpessoal e pode desaparecer em ambientes específicos ou determinadas situações (elevadores, grande quantidade de pessoas em um espaço pequeno). (HALL, 1989) propôs quatro tipos de distâncias interpessoais principais conforme pode ser observado no capítulo 2. Como o polígono do Voronoi de um dado sítio representa o conjunto de pontos mais próximos desse sítio do que qualquer outro, a utilização do VD como uma quantificação do espaço pessoal parece adequada.

Entretanto, conforme mencionado anteriormente neste trabalho, uma pessoa geralmente não percebe o espaço atrás dela. Tal fato inspirou a definição do o espaço pessoal percebido (PPS), que é a intersecção do campo de visão da pessoa (assumido como um setor circular cujo centro é dado pelo vetor de movimento da pessoa) e o seu espaço pessoal (ou seja, o seu polígono de Voronoi) (JACQUES et al., 2007). Uma ilustração gráfica do PPS pode ser vista na Figura 14.

Ainda conforme (JACQUES et al., 2007), o PPS pode ser explorado para caracterização do conforto de uma pessoa:

- **Pessoas confortáveis** - uma pessoa é dita confortável no grupo se a mesma possui um PPS (espaço pessoal percebido) relativamente grande, respeitando os limites de espaço definidos anteriormente por (HALL, 1989).
- **Pessoas desconfortáveis** - ao contrário de uma pessoa confortável, a pessoa des-

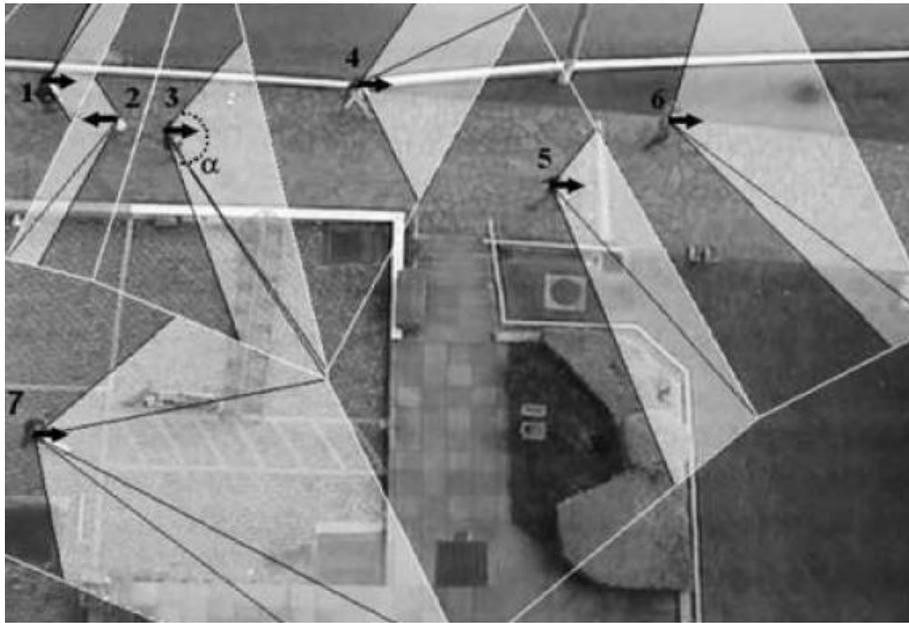


Figura 14: Diagramas de voronoi e espaço pessoal percebido (PPS)

confortável possui um PPS pequeno ¹.

Além disso, o conceito de conforto pode ser utilizado na classificação de grupos:

- **Grupo voluntário** - é um grupo cujos indivíduos que mantêm uma distância íntima durante um determinado período de tempo, com o PPS (espaço pessoal percebido) suficientemente grande. Isto significa que, embora os indivíduos percebam bastante espaço em seus campos de visão, eles escolheram permanecer próximos de outros indivíduos.
- **Grupo involuntário** - é um grupo dos indivíduos que mantêm uma distância íntima durante um determinado período de tempo, mantendo o PPS pequeno. Isto indica geralmente que eles formam um grupo porque não tem nenhum espaço livre para seguir.

Na primeira fase do modelo, as informações de entrada correspondem a extração de duas informações principais: definição do nível de conforto das pessoas na cena e classificação do grupo que formam, segundo o conforto extraído individualmente. As pessoas

¹Significando que provavelmente existem outras pessoas próximas e no seu campo de visão.

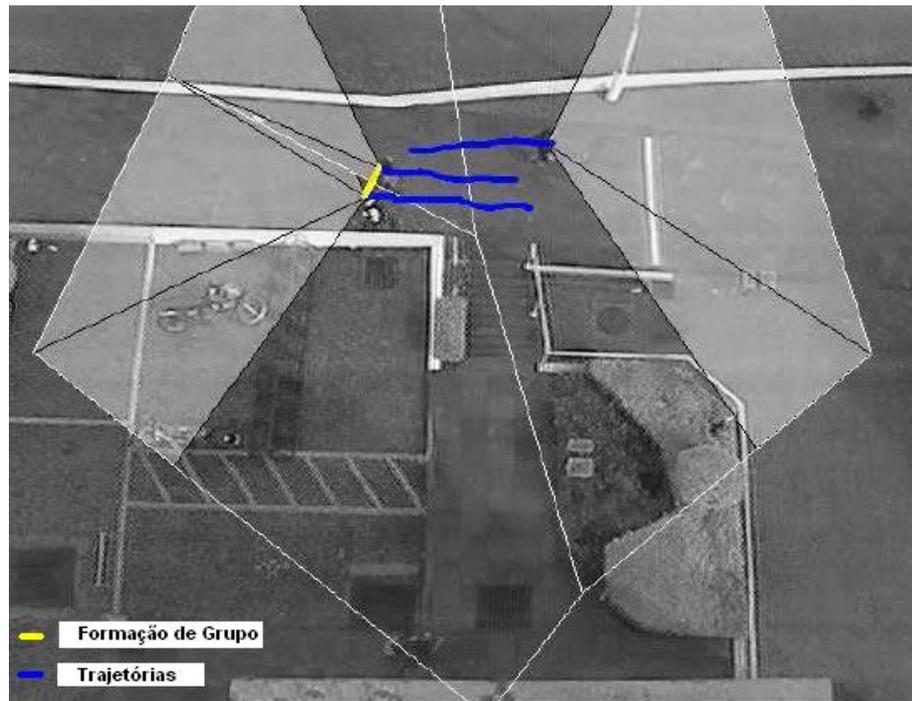


Figura 15: Detecção e classificação de grupos segundo (JACQUES et al., 2007)

detectadas na cena poderão ser classificadas como **confortáveis** ou **desconfortáveis** e o grupo em que elas pertencem, poderá ser definido como **voluntário** ou **involuntário**, dependendo das características do mesmos, conforme as definições acima. A Figura 15 mostra um exemplo de detecção e classificação de grupo.

Na Figura 15, é possível observar os campos de visão das pessoas segundo os diagramas de voronoi de cada uma, bem como as trajetórias das mesmas (linhas azuis). Neste exemplo, as pessoas foram detectadas na cena (3 pessoas) e um grupo foi detectado e classificado (duas pessoas - linha amarela). As pessoas presentes neste grupo estão num estado denominado **confortável** devido ao PPS ser relativamente grande nas duas pessoas e o grupo, desta forma é classificado como **voluntário**, pois estão numa distância curta uma pessoa da outra, mesmo tendo bastante espaço a sua disposição.

O algoritmo proposto por (JACQUES et al., 2007) visa classificar as pessoas que estão em movimento na cena. De fato, o algoritmo de acompanhamento de pessoas utilizado em (JACQUES et al., 2007) não consegue extrair a orientação de pessoas em estado estacionário (sem movimento), comprometendo o cálculo do PPS. Para contornar essa limitação e

permitir a análise de grupos estáticos, este trabalho propõe técnicas interativas para resolver o problema de orientação das pessoas paradas na cena. O método para resolução deste problema foi implementado da seguinte forma:

- o acompanhamento é executado para identificar as trajetórias, bem como as posições x e y de cada indivíduo durante o tempo corrente da cena;
- tendo em mãos as posições de cada indivíduo, é realizada uma verificação para saber se a variação da posição (x, y) daquela pessoa nos últimos N quadros foi maior que um certo limiar T . Caso seja maior ou igual, uma variável com valor 1 é atribuída a esta pessoa (pessoa em movimento); caso contrário, a variável com valor 0 é atribuída (pessoa em estado estacionário). Experimentalmente, notou-se que $N = 10$ e $T = 5$ se mostraram valores adequados para os vídeos empregados. Tais limiares devem ser alterados se a resolução dos vídeos for modificada (alteração em T), ou se a taxa de aquisição dos vídeos mudar (alteração em N).
- uma vez sabendo-se os quadros onde as pessoas estão em um estado estacionário, basta identificar os instantes de tempo em que esta situação ocorreu durante toda a cena, e informar o algoritmo responsável por processar os dados vindos do *tracking*;
- os dados informados ao *tracking* para pessoas que estão em estados estacionários são gerados de forma manual, através de cliques do mouse identificando a direção das pessoas.

Desta forma, tanto as pessoas em movimento ou em estado estacionário mantêm seus vetores de orientação de maneira correta, seja pelos dados gerados do *tracking* (para pessoas em movimento), ou pelos dados informados manualmente (para pessoas em estado estacionário).

Outro fator importante a ser citado é com relação a classificação dos grupos como voluntários ou involuntários. Quando o grupo detectado for um grupo em movimento, mantêm-se as definições propostas por (JACQUES et al., 2007) ou seja, se mais de 50% das

peças que compõem o grupo estiverem num estado confortável, o grupo é classificado como voluntário, caso contrário, é classificado como involuntário.

Porém, ao trabalhar-se com grupos sem movimento (estacionários), a classificação muda completamente. Grupos estáticos tendem a envolver pessoas interagindo (conversando) e então os membros normalmente ficam de frente uns para os outros. Assim, os PPSs são pequenos, e o grupo seria classificado como involuntário segundo (JACQUES et al., 2007). Este trabalho propõe uma nova classificação para grupos estáticos, classificando-os como voluntário se pelo menos 50% dos membros estiver desconfortável no grupo (ou seja, tiverem PPS pequeno), assim como contribuição para este trabalho é proposto um modelo de classificação de grupos estáticos (ou sem movimento) servindo de complemento ao modelo desenvolvido por (JACQUES et al., 2007) para grupos em movimento.

3.2 Fase 2: Estrutura do Grupo

Na segunda fase do modelo, as informações extraídas visam classificar a estrutura geométrica do grupo, ou seja, informações referentes ao polígono formado por este grupo, área e a variação desta no decorrer do tempo. A Figura 16 ilustra alguns possíveis casos:

1. cada pessoa detectada na cena é definida como um ponto, dando origem a uma forma geométrica distinta para os grupos ²;
2. uma questão a ser levada em conta refere-se à área que a estrutura geométrica possui. Por exemplo, uma grande área de um grupo pode trazer informações que as pessoas deste grupo estão mais dispersas, enquanto áreas pequenas podem representar maior envolvimento das pessoas na conversação;
3. algo a ser acompanhado ainda na estrutura refere-se à variação observada na área

²Se o grupo for composto por duas pessoas, a formação geométrica será uma linha, caso o grupo seja formado por três pessoas a geometria resultante será um triângulo, que poderá neste caso, ter subdivisões (triângulo equilátero, escaleno ou isósceles), dependendo do valor das distâncias entre os indivíduos do grupo.



Figura 16: Extração da estrutura do grupo segundo características pré-definidas

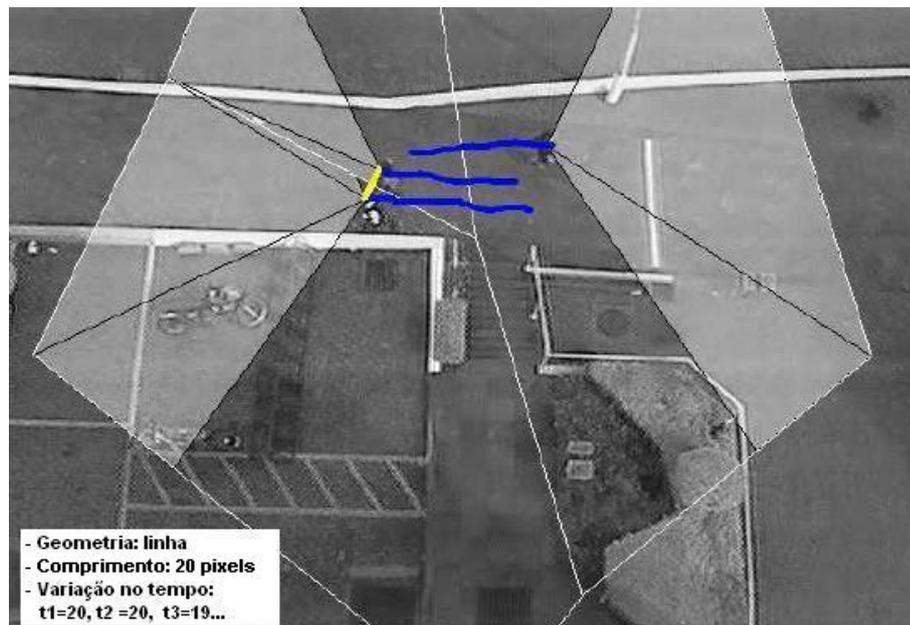


Figura 17: Extração de características segundo a estrutura do grupo

do polígono em função do tempo. É através da variação na área do polígono que se pode observar por exemplo, em quais períodos de tempo um grupo se mantém mais ou menos coeso, já que a coesão do grupo pode mudar dinamicamente.

Na fase 2, segundo o exemplo da Figura 17, as informações extraídas são referentes a estrutura do grupo, as quais são:

- geometria formada - linha (duas pessoas);
- comprimento - 20 *pixels*³;

³Nestes casos onde os valores são medidos em pixels, é necessário que o posicionamento da câmera bem como a altura da mesma sejam os mesmos para todos os exemplos a serem utilizados, desta forma sendo possível comparar os resultados extraídos de diferentes seqüências de vídeos

- comprimento em função do tempo: $t_1 \rightarrow 20$, $t_2 \rightarrow 20$, $t_3 \rightarrow 19$, etc.

Através destas três características principais a estrutura geométrica de um grupo é definida, e seus dados relativos a cada variável são armazenados e observados em função do tempo.

3.3 Fase 3: Atividade do Grupo

A atividade de um grupo é a terceira etapa deste modelo e visa reunir características para saber o quanto um grupo (ou seus membros individualmente) é ativo (pessoas conversando, gesticulando) ou inativo (pessoas sem movimentação/interação). Para saber estas informações, o primeiro passo é medir as interações presentes entre os indivíduos que compõem o grupo.

Segundo (SILVA, 2007), o termo *interação* é utilizado para qualificar qualquer coisa cujo funcionamento permite algum nível de participação ou troca de ações. Ainda segundo (SILVA, 2007) o conceito de interação social começou a ser utilizado a partir do início do século XX, e designa a influência recíproca dos atos de pessoas ou grupos.

Para medir as interações presentes entre os indivíduos que fazem parte do grupo, assume-se que as relação entre pares de pessoas acontecem quando elas estão aproximadamente de frente uma para a outra. Assim, deve-se capturar a orientação destes ⁴, e avaliar a relação entre as orientações e posições relativas de cada pessoa e seus vizinhos. Sendo \mathbf{x}_i a posição de cada pessoa na cena, e sendo \mathbf{v}_i as respectivas orientações, calcula-se o quanto a pessoa i está “visando” a pessoa j através de:

$$I_{ij} = \frac{\langle \mathbf{v}_i, (\mathbf{x}_j - \mathbf{x}_i) \rangle}{\|\mathbf{v}_i\| \|\mathbf{x}_j - \mathbf{x}_i\|}. \quad (3.1)$$

⁴Conforme já mencionado no trabalho, a orientação é obtida através da trajetória para pessoas em movimento, e fornecida manualmente para pessoas em estado estacionário (sem movimento).

Analogamente, calcula-se o quanto a pessoa j está “visando” a pessoa i através de:

$$I_{ji} = \frac{\langle \mathbf{v}_j, (\mathbf{x}_i - \mathbf{x}_j) \rangle}{\|\mathbf{v}_j\| \|\mathbf{x}_i - \mathbf{x}_j\|}. \quad (3.2)$$

As Equações (3.1) e (3.2) representam o produto interno entre os vetores normalizados que caracterizam a orientação das pessoas e os vetores entre eles. Se a pessoa i estiver olhando diretamente para a pessoa j , então $I_{ij} = 1$. Se a pessoa i estiver de costas para a pessoa j , então $I_{ij} = -1$. O mesmo raciocínio vale para a relação entre a pessoa j e a pessoa i .

A interação máxima é obtida quando as duas pessoas estão frente à frente, e a mínima quando alguma está de costas para a outra. Assim, a interação conjunta $I'_{i,j}$ entre as pessoas i e j pode ser quantificada via:

$$I'_{i,j} = \min\{I_{ij}, I_{ji}\}. \quad (3.3)$$

Também é proposta uma classificação da interação entre pessoas em três níveis:

- Interação fraca: ocorre se $I'_{i,j} < I_{\min}$;
- Interação mediana: ocorre se $I_{\min} < I'_{i,j} < I_{\max}$;
- Interação forte: ocorre se $I'_{i,j} > I_{\max}$.

Neste trabalho, foi determinado empiricamente que $I_{\min} = 0.26$ e $I_{\max} = 0.87$, que correspondem a ângulos de aproximadamente 75° e 30° , respectivamente.

Um outro ponto fundamental nesta fase corresponde à verificação da variação dos *blobs* dos indivíduos que compõem o grupo. Em grupos com interação, as pessoas mais ativas (quem está falando, por exemplo) tendem a gesticular e se movimentar mais do que as passivas (ouvintes). Para grupos estáticos, acredita-se que tal atividade pode ser avaliada acompanhando a variação temporal da área do *blob*⁵. Assim, a análise dos

⁵Já para grupos andando, o movimento dos braços e pernas causam uma variação dinâmica na área dos *blobs*, e assim o movimento de gesticular provavelmente será ofuscado pelo movimento de caminhada.

blobs individuais pode trazer informações como a movimentação dos agentes no grupo, importante no que tange a atividade dos grupos.

Nesta etapa, é importante considerar que pessoas muito próximas umas das outras podem gerar um *blob* (neste trabalho, chamado de *blob* fundido). Dada uma seqüência de vídeo, um limiar T_a é estipulado para denotar a maior área de *blob* esperada na subtração de fundo. Assim, *blobs* fundidos são detectados como aqueles que possuem área maior do que T_a .

O cálculo da variação da área é direto. Sendo $A_i(t)$ a área da pessoa i no quadro t , sua variação instantânea $\delta A(t)$ no quadro t é dada por

$$\delta A(t) = A(t) - A(t - 1), \quad (3.4)$$

e variação média da área entre os quadros t_0 e t_f é dada por

$$\delta A_m(t_0, t_f) = \frac{1}{t_f - t_0 + 1} \sum_{t=t_0}^{t_f} |\delta A(t)|. \quad (3.5)$$

Neste trabalho, assume-se que valores altos para $\delta A_m(t_0, t_f)$ correspondem a pessoas com uma maior atividade em grupos estáticos.

Outro critério que pode ser extraído diretamente das variações instantâneas é a variação acumulada $\delta A_c(t, m)$, que representa um somatório das variações instantâneas nos últimos m quadros antes do quadro t :

$$\delta A_c(t, m) = \sum_{k=t-m}^t \delta A(k) = A(t) - A(t - m - 1), \quad (3.6)$$

que representa a diferença de área calculada entre os quadros t e $t - m - 1$.

Caso uma mudança na estrutura do grupo seja detectada pelo sistema, as fases 1, 2 e 3 devem ser acionadas novamente, para que os novos padrões sejam aplicados ao novo grupo detectado. Um grupo pode sofrer mudanças durante um determinado período de tempo, seja por elementos que saem ou entram em um grupo, pelo término

do grupo, etc. Torna-se assim necessária a redefinição tanto da classificação deste novo grupo bem como do processo de extração de características deste grupo, para que seja possível redefini-lo dentro de um novo padrão. Assim, tem-se todas as fases necessárias para criação e execução do modelo proposto. O capítulo a seguir mostra em detalhes o protótipo desenvolvido.

Capítulo 4

Protótipo

A implementação do sistema foi desenvolvida através do *software* Matlab, desde os ajustes dos parâmetros iniciais até a geração dos gráficos finais. Os arquivos principais do sistema podem ser visualizados na Figura 18.

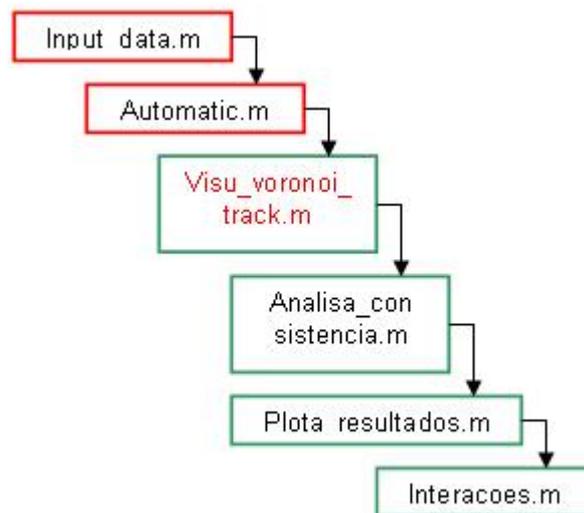


Figura 18: Estruturação dos principais arquivos utilizados na modelagem do sistema

O arquivo *input-data.m* é o arquivo de configuração dos parâmetros iniciais do sistema. É neste arquivo que define-se o local das imagens a serem processadas, o formato destas imagens, a imagem inicial e final para treinamento de *background* e execução do *tracking*, definição da área para detecção do *blob* da pessoa na cena, além de outras

informações quanto a configuração do *tracking* a ser processado.

Uma vez definida a estrutura do arquivo inicial (*input-data.m*), o arquivo *automatic.m* é executado. Este arquivo tem como finalidade detectar de forma automática as pessoas que estão na cena e acompanhá-las durante a execução da mesma. Além disso, o arquivo *automatic.m* separa os objetos de *background* (sombras, obstáculos, etc.), dos objetos de *foreground* (neste caso as pessoas da cena). Um exemplo de aplicação deste arquivo pode ser visto na Figura 19.

No quadro (A) os pontos vermelhos referem-se as pessoas que já foram detectadas na cena como *foreground* e que estão sendo acompanhadas, considerando para isto um tempo mínimo existente de quadros, para o objeto ser considerado como *foreground*. No quadro (B) tem-se as pessoas da cena, com a detecção da sombra das mesmas (em azul). No quadro (C) ocorre a separação da sombra (*background*) com a pessoa da cena. O quadro (D) mostra a imagem de *background* onde ocorre a cena definida ainda no arquivo *input-data.m*, no período de treinamento. No quadro (E) tem-se os *blobs* individuais das pessoas detectadas na cena, podendo fundir-se dependendo da proximidade dos mesmos. O quadro (F) mostra a cena (imagem) real que está sendo processada, com a respectiva área de cada pessoa (quadrado verde) e a detecção da estimativa da cabeça desta pessoa na cena (quadrado vermelho).

Realizado o *tracking*, o arquivo *visu-voronoi-track.m* é executado. Este arquivo tem como objetivo principal receber os dados do arquivo *automatic.m*, que contém as posições (x,y) das pessoas detectadas na cena e aplicar o conceito de espaço pessoal percebido juntamente com os diagramas de voronoi.

Além disso, o arquivo *visu-voronoi-track.m*, possui as principais funções do sistema, onde é possível estimar as distâncias entre as pessoas, segundo o conceito definido por (HALL, 1989), detectar e classificar o grupo como voluntário ou involuntário, através do nível de conforto das pessoas do grupo, além da captura da trajetória das pessoas, tempo da cena, entre outras informações. Um exemplo deste arquivo em execução pode ser visto

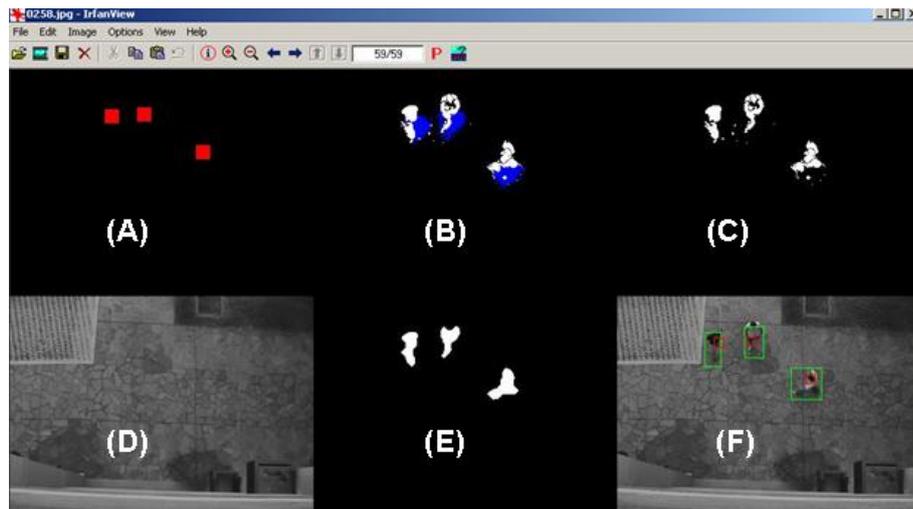


Figura 19: Exemplo de aplicação do arquivo *automatic.m*

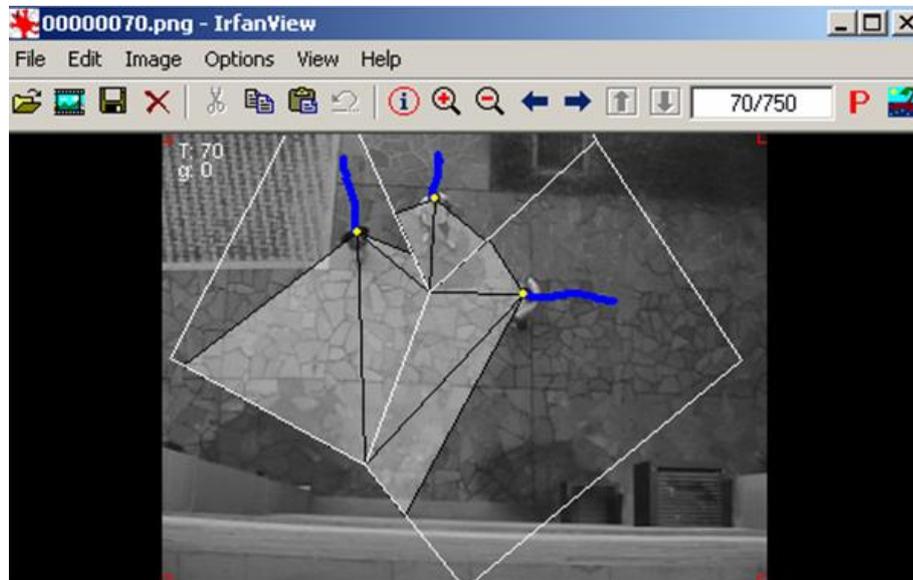


Figura 20: Imagem extraída na execução do arquivo *visu-voronoi-track.m*

conforme a Figura 20.

O arquivo *analisa-consistencia-do-grupo.m* é responsável por gerar/processar todos os dados referentes ao grupo que está sendo analisado. Com isto tem-se um vetor chamado de **Identificacao-do-grupo** que armazena todas as informações que deseja-se extrair do grupo na cena.

O arquivo *analisa-consistencia-do-grupo.m* é uma função, chamada dentro do arquivo *visu-voronoi-track.m*. Existem ainda sub-funções dentro deste arquivo que fazem parte do cálculo dos dados que estão sendo armazenados dentro do vetor **Identificacao-**

Vetor - IDENTIFICACAO_DO_GRUPO			
Grupos de 2 pessoas		Grupos de 3 pessoas	
1	Tempo (Frame)	1	Tempo (Frame)
2	Número de Pessoas	2	Número de Pessoas
3	Distância	3	Area
4	Geometria formada	4	Distância 1
5	Centróide x	5	Distância 2
6	Centróide y	6	Distância 3
7	Classificação do Grupo (Voluntário / Involuntário)	7	Geometria formada
8	Posição x1 (pessoa 1)	8	Centróide x
9	Posição y1 (pessoa 1)	9	Centróide y
10	Blob da pessoa 1	10	Classificação do Grupo (Voluntário / Involuntário)
11	Posição x2 (pessoa 2)	11	Posição x1 (pessoa 1)
12	Posição y2 (pessoa 2)	12	Posição y1 (pessoa 1)
13	Blob da pessoa 2	13	Blob da pessoa 1
		14	Posição x2 (pessoa 2)
		15	Posição y2 (pessoa 2)
		16	Blob da pessoa 2
		17	Posição x3 (pessoa 3)
		18	Posição y3 (pessoa 3)
		19	Blob da pessoa 3

Figura 21: Vetor - Identificação do Grupo - grupos formados por duas e três pessoas

do-grupo. As informações que estão sendo gravadas dentro desta lista dividem-se em grupos formados por duas ou três pessoas, e podem ser observadas conforme mostra a Figura 21. Analisando as informações do vetor conforme a Figura 21, tem-se no primeiro campo o armazenamento do **tempo** na cena, que armazena o quadro onde começa a detecção de um grupo, informação passada através da variável *TimeStep* armazenada no arquivo *visu-voronoï-track.m*.

Em seguida, é armazenado o **número de pessoas** que compõem o grupo detectado, através da variável *grupo-count* que é ativada caso um grupo seja identificado pelo sistema. A informação quanto ao número de pessoas presentes no grupo é fundamental, pois dependendo do valor retornado, são ativadas as variáveis para armazenar os valores correspondentes aos grupos formados por duas ou três pessoas. Por exemplo, se o valor retornado é dois (grupo de duas pessoas), então um total de treze informações terão que ser

armazenadas no vetor **Identificação-do-grupo**, caso o valor retornado seja três (grupo de três pessoas), um total de dezenove informações serão armazenadas no vetor.

A informação de **distância** é medida quando o grupo possui duas ou três pessoas, utilizando-se para isto, a fórmula da distância entre dois pontos, para o cálculo entre as duas pessoas, extraíndo suas posições x e y que estão armazenadas numa matriz de posições das pessoas da cena.

Para cálculo da **área do grupo**, que é ativada no sistema quando o grupo é composto por três pessoas, utiliza-se o cálculo da área do triângulo.

Quanto a **geometria** formada pelo grupo, pode-se diferenciar em dois aspectos: quando o grupo for composto por duas pessoas a geometria padrão será uma linha, identificado no sistema pelo número 99; quando se tratar de três pessoas no grupo, é verificado se as três distâncias entre as pessoas são iguais ¹, neste caso forma-se um triângulo equilátero e recebe-se o valor 111 (valor simbólico para representar a geometria triângulo equilátero, da mesma forma ocorre para os valores referentes a triângulo isósceles - 122 e triângulo escaleno - 123). Caso duas distâncias sejam iguais e uma diferente, corresponde a um triângulo isósceles, o valor recebido será 122, e no último caso, caso as três distâncias sejam diferentes, forma-se um triângulo escaleno, recebendo o valor 123.

O **centróide** do grupo tem por objetivo armazenar os valores do centro do grupo (média das posições (x,y) das pessoas do grupo), para saber se o grupo em questão, seja grupos de duas ou três pessoas, está em movimento (andando, correndo, etc.) ou sem movimento.

A **classificação do grupo** é diretamente ligada ao conforto das pessoas que compõem o grupo (para grupos em movimento). Se mais de 50% das pessoas que formam o grupo estiverem num estado denominado confortável (áreas de visão não obstruídas, PPS grande) então o grupo é classificado como **voluntário** recebendo no vetor **Identi-**

¹No caso da geometria ser um triângulo, duas distâncias são consideradas iguais verificando uma igualdade aproximada. De fato, elas são consideradas iguais se sua diferença (em módulo) for menor ou igual do que um limiar Δx . Neste trabalho, dada a resolução das imagens capturadas, foi estipulado um limiar $\Delta x = 2 \text{ pixels}$.

ficação-do-grupo valor 555, caso contrário, se o PPS das pessoas que formam o grupo for pequeno, o grupo é classificado como **involuntário**, recebendo o valor de 666. Para grupos sem movimento ou parados, a lógica se inverte, sendo que os grupos podem ter um PPS pequeno e serem classificados como voluntários, caso as pessoas que compõe o grupo estejam com seus vetores direção apontando uma para outra (no caso, interagindo).

As **posições das pessoas** na cena são armazenadas em matrizes com a posição (x,y) de cada uma, ainda no arquivo *automatic.m*. O arquivo *analisa-consistencia-do-grupo.m* busca estes valores, através da passagem de parâmetros entre as funções que estão ligadas a estes dois arquivos.

Os **blobs individuais** das pessoas do grupo são extraídos de forma automática, ainda no arquivo *automatic.m*. A extração desta informação irá revelar se a área das pessoas ativas em um grupo varia mais do que as pessoas que estão estáticas naquele grupo, ao longo do tempo. Porém para grupos em movimentos esta informação torna-se difícil de analisar, pois as pessoas se “mexem” ao caminhar.

4.1 Mostrando resultados graficamente

Para que o modelo proposto pudesse disponibilizar a comparação entre as variáveis do sistema, bem como ter uma forma de representação dos dados de maneira visual (através de gráficos), foi necessário o desenvolvimento de um arquivo denominado *plota-resultados.m*. Este arquivo tem como função armazenar os valores de cada tempo da cena, separados em células, para que seja possível manipular os dados de forma correta e organizada.

Desta forma, é possível comparar qualquer variável que está armazenada no vetor **Identificacao-do-grupo** para analisar se os dados gerados são coerentes com o vídeo e se as informações estão corretas. A Figura 22 mostra uma seqüência de quadros onde ocorre formação de um grupo de três pessoas, sendo que uma delas se afasta em alguns

períodos de tempo, resultando em um grupo de duas pessoas temporariamente. O gráfico comparativo entre o número de pessoas presentes no grupo durante o decorrer da cena pode ser visualizado conforme mostra a Figura 23. No gráfico da Figura 23 é realizada uma comparação entre o *tempo* (horizontal), correspondente ao tempo de vida do grupo formado, com o número de pessoas existentes no grupo (vertical), sendo possível visualizar graficamente, em que determinada faixa de tempo uma formação de grupo começou e terminou e também, quantos eventos foram detectados pelo sistema (neste caso, quando houve uma mudança na formação de grupos de duas pessoas para três pessoas ou vice-versa).

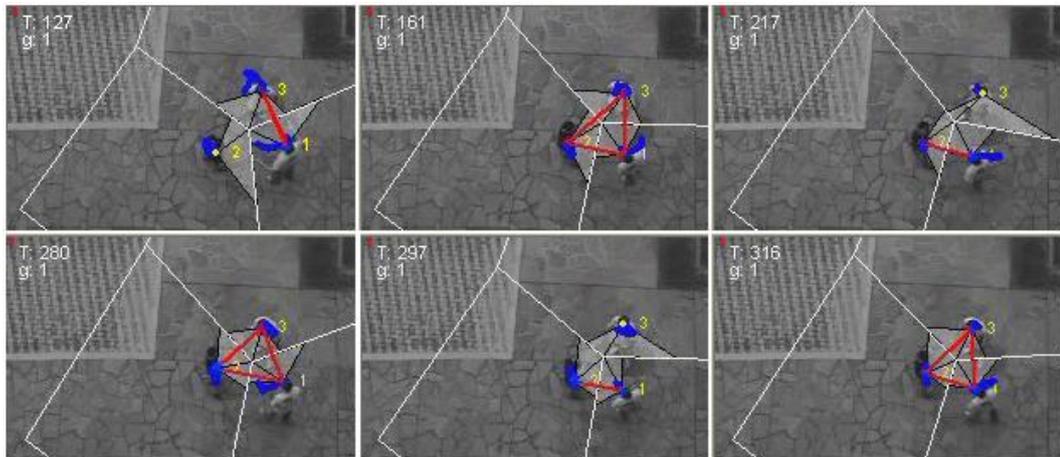


Figura 22: Seqüência de agrupamento formado por duas e três pessoas durante amostragem de vídeo

4.2 Calculando as interações do grupo

Uma parte importante na extração de dados de grupo diz respeito à interação entre as pessoas da cena. A interação visa avaliar a orientação de cada pessoa e verificar qual(is) pessoa(s) está(ão) interagindo durante um determinado tempo, valendo para grupos formados por duas ou três pessoas. Conforme descrito na Seção 3.3, as interações entre pares de pessoas são calculada com base nos produtos internos normalizados dos vetores orientação e vetores relativos entre as pessoas.

Ao final da execução, o arquivo *interações.m* salva as imagens quadro a quadro,

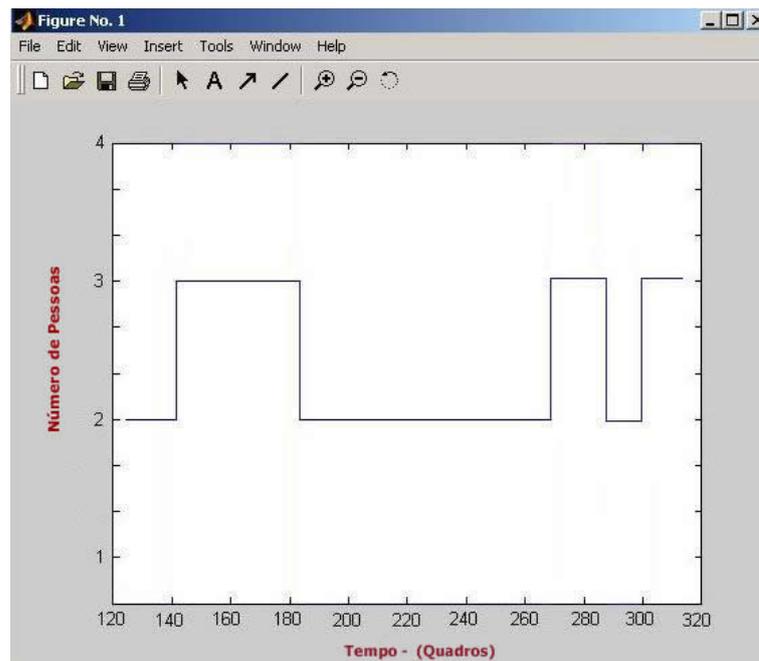


Figura 23: Visualizando os resultados graficamente

mostrando visualmente através de uma linha vermelha quais foram as interações, entre quais indivíduos do grupo e em quais quadros eles ocorreram. Um exemplo disso pode ser visto na Figura 24.



Figura 24: Detecção da interação entre as pessoas do grupo

4.3 Interface WEB para classificação das informações

Uma interface web foi desenvolvida com o objetivo de classificar as informações geradas pelo protótipo em Matlab, tornando possível a atualização dinâmica das informações que são armazenadas no banco de dados. Esta ferramenta foi projetada para complemen-

tar o protótipo desenvolvido, tendo em vista que o mesmo gera os resultados em arquivos no formato texto (vetor IDENTIFICAÇÃO DO GRUPO e também o arquivo referente as interações entre os indivíduos do grupo), fazendo-se necessário o desenvolvimento de uma ferramenta que complementa-se à parte gráfica e que pudesse gerar resultados dinamicamente para o usuário, através da inserção das informações no banco de dados.

A interface web foi desenvolvida através da linguagem de programação PHP (devido a atualização dinâmica dos dados), utilizando-se o banco de dados MySQL (para armazenar/inserir os valores extraídos do vetor **Identificacao-do-grupo**) e o servidor web Apache (para hospedar a aplicação). A tela principal (index.php) pode ser visualizada conforme mostra a Figura 25.



Figura 25: Tela principal da interface web

As principais funcionalidades da interface são:

- seleção de trechos de vídeo: são disponibilizados ao usuário diversos vídeos onde é possível escolher de qual deles deseja-se extrair informações;
- resultados processados: esta opção traz um resumo das informações do vídeo,

mostrando as informações em percentuais e gerando dados para a classificação final do grupo;

- **classificação final:** esta opção mostra para o usuário a classificação final do grupo, com relação ao trecho de vídeo selecionado.

A interface web desenvolvida é de fácil utilização e intuitiva para o usuário. Conforme a ilustração da Figura 25, ao clicar sobre uma das imagens que se encontram no centro da tela principal (em formato reduzido), o vídeo correspondente é carregado na tela e o usuário pode verificar se é realmente este trecho de vídeo que ele quer analisar. Vale lembrar que este protótipo poderia servir como aplicação futura para sistemas de monitoramento de vídeo ou vigilância, onde câmeras espalhadas por diferentes lugares geram inúmeros vídeos. Selecionado o vídeo, o usuário poderá ver o conteúdo do processamento dos dados extraídos pelo protótipo em Matlab, de forma detalhada e completa conforme mostra a Figura 26.

Os itens da Figura 26 mostram em sua ordem:

- **frame (duração do agrupamento)** - corresponde ao tempo total (em quadros) que durou o agrupamento dos indivíduos (a taxa de aquisição dos vídeos foram realizados à 30 FPS);
- **número de pessoas** - indica durante quanto tempo o grupo foi formado por duas ou três pessoas, juntamente com seus percentuais;
- **distância entre pessoa 1 e 2** - informa a distância mínima, média e máxima mantida durante a formação do grupo, com relação a pessoa 1 e 2 (grupos de dois indivíduos);
- **área** - indica o tamanho mínimo, médio e máximo da área formada pelo grupo de três pessoas, durante o decorrer da cena;

- **distância entre as pessoas do grupo** - para o grupo formado por três indivíduos, as distâncias mínimas, médias e máximas são calculadas entre as pessoas 1 e 2, 2 e 3 e por fim entre as pessoas 3 e 1;
- **geometria** - tem o objetivo de informar a geometria formada pelo grupo durante a formação do mesmo. O padrão *linha* é para grupos formados por duas pessoas, *triângulo equilátero*, para o grupo de três pessoas onde as distâncias entre os membros foram todas iguais, *triângulo isósceles*, para grupos onde duas distâncias entre os indivíduos são iguais e uma distância é diferente, *triângulo escaleno*, para grupos onde os indivíduos possuem três distâncias diferentes entre si;
- **centróide (x, y)** - corresponde ao centróide formado pelo grupo nas coordenadas (x, y) , mostrando assim a posição mínima, média e máxima nas duas coordenadas, onde o grupo formou-se.
- **classificação do grupo: voluntário ou involuntário** - mostra por quanto tempo o grupo formado foi voluntário ou involuntário, conforme o conceito de grupos voluntários e involuntários definidos no modelo proposto;
- **posição da pessoa em (x, y)** - mostra a posição da pessoa durante a cena (mínima, média e máxima) nas coordenadas (x, y) ;
- **blob da pessoa** - mostra a área do *blob* formada pela pessoa (mínima, média e máxima) durante o decorrer da cena. Além disso, é informada a quantidade e porcentagem de *blobs* normais, ou seja, que não se fundiram com outros *blobs* e também a quantidade e porcentagem de *blobs* fundidos.
- **interação entre as pessoas do grupo** - a interação mínima, média e máxima entre os membros do grupo é exibida, levando em consideração o decorrer da cena. Além disso, foram criados três critérios para avaliar a interação entre os indivíduos do grupo: **sem interação** - são as pessoas que não tiveram nenhuma interação durante o decorrer da cena; **interação média ou transitória** - foram os indivíduos que

tiveram algum tipo de interação com outro indivíduo; **com interação** - corresponde aos indivíduos que tiveram interação com outro indivíduo devido aos vetores direção estarem totalmente voltados para outra pessoa, indicando que houve interação entre os mesmos. Por fim a porcentagem das interações é mostrada.

Com relação a classificação final gerada para os grupos, foram levadas em consideração as maiores porcentagens dos principais itens, tais como:

- número de pessoas;
- geometria formada;
- classificação do grupo: voluntário ou involuntário;
- variação média das áreas dos *blobs* das pessoas que compõem o grupo (conforme Equação (3.5));
- porcentagem de interação mediana e alta em relação as pessoas que formam o grupo.

Uma melhor representação dos conceitos aqui expostos, bem como os estudos de caso que contém as variáveis apresentadas neste capítulo podem ser observados no capítulo 5, referente aos resultados obtidos.

Classificação de grupos utilizando informações de geometria e detecção de atividades intragrupo									
Vídeo selecionado: Vídeo 1 									
	Total								
Frame:	58								
	Duas	Três	Porc. 2	Porc. 3					
Número de Pessoas:	24	34	41.38 %	58.62 %					
	Mínima	Média	Máxima						
Distância entre Pessoa 1 e 2: (Grupo de 2 pessoas)	26.01	36.2	46.17						
	Mínima	Média	Máxima						
Area:	337.5	454.74	657						
	Mínima	Média	Máxima						
Distância entre Pessoa 1 a 2: (Grupo de 3 pessoas)	25	29.35	36						
Distância entre Pessoa 2 a 3: (Grupo de 3 pessoas)	31	38.18	47						
Distância entre Pessoa 3 a 1: (Grupo de 3 pessoas)	28	31.59	37						
	Linha	Triangulo Equilatero	Triangulo Isosceles	Triangulo Escaleno	Porc. L	Porc. TE	Porc. TI	Porc. TES	
Geometria:	24	0	1	33	41.38 %	0 %	1.72 %	56.9 %	
	Mínima	Média	Máxima						
Centróide de X:	136.5	140.98	145.33						
Centróide de Y:	72.33	79.18	88						
	Quant. Volunta.	Quant. Involunta.	Porc. Volunta.	Porc. Involunt.					
Classificação Voluntario/Inv.:	58	0	100%	0%					
	Mínima	Média	Máxima						
Posição da Pessoa x1:	150	155.64	166						
Posição da Pessoa y1:	80	84.64	89	Quant. Normais	Quant. Fundidos	Porc. Normais	Porc. Fundidos	Variação média da área	
Blob da Pessoa 1:	327	653.97	1424	35	23	60.34%	39.66%	40,75	
Posição da Pessoa x2:	118	123.75	127						
Posição da Pessoa y2:	83	85.27	88	Quant. Normais	Quant. Fundidos	Porc. Normais	Porc. Fundidos	Variação média da área	
Blob da Pessoa 2:	278	427.48	1473	42	2	95.45%	4.55%	30,20	
Posição da Pessoa x3:	120	138.98	155						
Posição da Pessoa y3:	48	64	89	Quant. Normais	Quant. Fundidos	Porc. Normais	Porc. Fundidos	Variação média da área	
Blob da Pessoa 3:	316	617.69	1424	34	14	70.83%	29.17%	34.51	
	Mínima	Média	Máxima	Sem interação	Interação mediana	Com interação	Porc. Sem	Porc. Media	Porc. Com
Interação entre Pessoa 1 e 2:	-0.07	-0.4	-0.98	0	21	23	0%	47.73%	52.27%
Interação entre Pessoa 2 e 3:	-0.18	-0.51	-0.75	0	10	24	0%	29.41%	70.59%
Interação entre Pessoa 3 e 1:	0.27	-0.52	-0.9	0	10	38	0%	20.83%	79.17%

Figura 26: Demonstração dos resultados extraídos do protótipo em Matlab através da interface web

Capítulo 5

Resultados Obtidos

Neste capítulo, são exibidos três estudos de caso com relação ao protótipo proposto, com o intuito de comprovar a eficácia do mesmo, bem como verificar a relação existente entre os dados extraídos do protótipo com os vídeos reais onde a detecção automática de grupos foi aplicada. As seqüências de vídeo utilizadas para os testes foram desenvolvidas nesta universidade (UNISINOS/RS), em frente ao prédio 6, junto ao PIPCA (Programa Interdisciplinar de Pós-Graduação em Computação Aplicada). O objetivo dos experimentos era formar grupos de duas e três pessoas respectivamente em situações de movimento e estado estacionário, além de outras características, visando detectar grupos, extrair as informações necessárias, gerar sua classificação automática e semi-automática (em alguns casos).

Vale lembrar que os dados extraídos pelo protótipo desenvolvido no software Matlab foram integrados com uma interface web, desenvolvida para mostrar os resultados graficamente de maneira detalhada, organizar de forma dinâmica os dados, bem como gerar a classificação final dos vídeos analisados. Esta ferramenta web pode servir para futuras aplicações onde este protótipo for utilizado, conforme necessidade.

Todos os estudos de caso (1, 2 e 3) foram realizados com a mesma distância e posicionamento da câmera, de forma que os dados extraídos de cada grupo pudessem ser

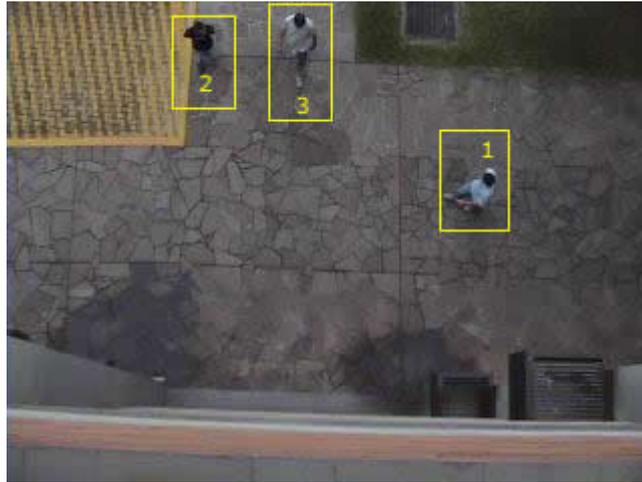


Figura 27: Caso A: formação de um grupo composto por três pessoas

comparados entre si. A taxa de aquisição de vídeo em todos os exemplos foi de 30 FPS (trinta quadros por segundo).

5.1 Caso A: Encontro de três amigos

Neste primeiro vídeo tem-se a formação de um grupo constituído por três indivíduos, onde dois indivíduos entram na cena no topo da imagem e um terceiro indivíduo entra pelo lado direito da cena (lado direito da imagem), conforme pode ser visto na Figura 27.

Cada indivíduo que entra na cena e é detectado, recebe um número de identificação (ID). Este ID é atribuído pela ordem de entrada na cena do indivíduo e serve para visualizar e mapear as pessoas que fazem parte do vídeo.

Uma vez detectadas as pessoas no vídeo, o protótipo entra em ação, para extrair os dados relativos ao grupo formado e demais características. Neste primeiro caso, os principais dados gerados foram os seguintes:

- tempo (duração do agrupamento): 183 quadros;
- número de pessoas que formam o grupo: duas pessoas = 55,91%, três pessoas =

44,09%;

- geometria formada: linha = 44,09%, triângulo isósceles = 0,79%, triângulo escaleno = 55,12%;
- classificação do grupo: voluntário = 100%, involuntário 0%;
- *blob* da pessoa 1: *blobs* normais = 77,17%, *blobs* fundidos = 22,83%, variação média da área do *blob* = 41,68 pixels;
- *blob* da pessoa 2: *blobs* normais = 97,35%, *blobs* fundidos = 2,65%, variação média da área do *blob* = 32,27 pixels;
- *blob* da pessoa 3: *blobs* normais = 74,42%, *blobs* fundidos = 25,48%, variação média da área do *blob* = 36,14 pixels;
- interação entre pessoa 1 e 2: sem interação = 20,35%, interação mediana = 77,88%, com interação = 1,77%;
- interação entre pessoa 2 e 3: sem interação = 43,89%, interação mediana = 55,01%, com interação = 1,10%;
- interação entre pessoa 3 e 1: sem interação = 36,22%, interação mediana = 62,20%, com interação = 1,58%.

Com relação ao segundo item citado (**número de pessoas que formam o grupo**) faz-se necessário esclarecer os dados apresentados acima. Um grupo é formado quando duas ou mais pessoas encontram-se em uma distância íntima conforme classificação de (HALL, 1989). Desta forma, estipula-se no algoritmo uma distância em pixels para cada uma das distâncias interpessoais classificadas por (HALL, 1989) (distância íntima, pessoal, social e pública), com base na posição e altura da câmera. Assim, as duas pessoas na cena que ficaram em uma distância íntima primeiramente foram as pessoas com ID 1 e 2, Figura 28, quadro A, permanecendo de tal modo durante 44,09% da cena. A terceira pessoa da cena (ID 3) aproximou-se e juntou-se ao grupo alguns quadros depois (quadro B), sendo

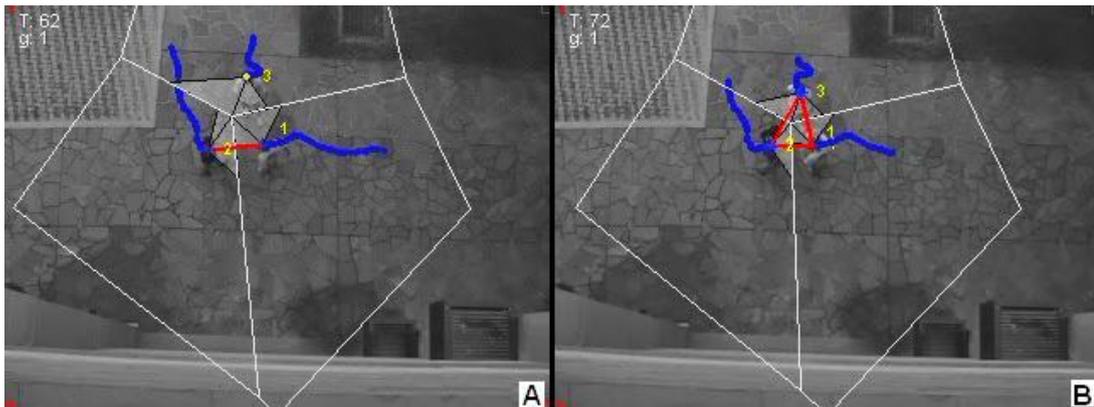


Figura 28: Formação de grupo envolvendo duas e três pessoas em distância íntima que o grupo foi formado por três pessoas durante 55,91% da cena. Esta representação pode ser visualizada na Figura 28.

O terceiro item traz informações quanto à geometria formada pelo grupo. Neste caso tem-se quatro alternativas: linha, triângulo equilátero, isósceles e escaleno (conforme mencionado anteriormente). Os dados do terceiro item do Caso A mostram que em 44,09% do tempo a geometria formada foi uma linha; 0,79% triângulo isósceles; 55,12% triângulo escaleno e nenhum triângulo equilátero foi detectado neste primeiro caso (desta forma é possível afirmar que na maior parte do vídeo as pessoas mantiveram distâncias diferentes entre si durante a formação do grupo).

No item quatro, tem-se a classificação do grupo como voluntário ou involuntário. Vale lembrar que para grupos em movimento a classificação utilizada segue o proposto no trabalho de (JACQUES et al., 2007), enquanto que para grupos em estado estacionário segue as definições conforme o capítulo 3, referente ao modelo proposto. Neste item quatro a classificação do grupo foi 100% do tempo um grupo voluntário, sendo que os indivíduos que compõem este grupo estiveram todo o tempo interagindo entre si.

Os itens 5, 6 e 7 referem-se aos *blobs* das pessoas 1, 2 e 3 respectivamente. O conceito de *blobs* individuais foi abordado e implementado neste trabalho, com o intuito de poder medir o nível de atividade do grupo, ou seja, identificar qual grupo tem maior movimentação entre seus membros, ou qual grupo está mais parado (sem movimentação).

Neste caso, a área do *blob* é verificada quadro-a-quadro para verificar sua variação no decorrer do tempo.

Os dados gerados no item 5 mostram os *blobs* normais, *blobs* fundidos e variação média do *blob* (medida em *pixels*). Os *blobs* normais correspondem a denominação dada para os *blobs* que não se juntaram a outro *blob* quando da execução do acompanhamento (*tracking*). Já os *blobs* fundidos foram os *blobs* que estavam muito próximos de outro *blob*, na qual geraram uma área muito alta devido a fusão de um *blob* com outro, sendo necessário desconsiderar a área quando detectada a fusão.

Algo interessante a ser ressaltado neste item é que, nesta pesquisa, um dos objetivos era saber se área do *blob* das pessoas que estavam conversando e/ou gesticulando em um grupo apresentava uma variação maior do que a variação do *blob* das pessoas que aparentemente estavam paradas durante a formação do grupo. De fato, esta característica comprovou-se experimentalmente, ao menos nos vídeos que foram analisados. As pessoas que gesticulam ou efetuam movimentos durante o diálogo em grupo, mantêm uma variação da área maior do que as pessoas que estão sem movimentar-se no grupo, ou pouco participativas. Neste estudo de caso A, a variação dos *blobs* das pessoas 1, 2 e 3 foi respectivamente de 41.68, 32.27 e 36.14 pixels, taxas altas se comparadas com a variação de outros grupos onde as pessoas movimentam-se pouco durante a formação do grupo.

A Figura 29 mostra um exemplo de variação da área dos *blobs* no decorrer do tempo. As pessoas 1 e 3 da ilustração indicam variação alta com relação a área dos *blobs*, já a pessoa 2 indica variação da área do *blob* baixa o que demonstra ser uma pessoa sem movimentação no grupo que pertence. De fato, no gráfico é possível visualizar que as informações extraídas pelo protótipo são coerentes.

Os itens 8, 9 e 10 tratam da interação entre as pessoas que compõem o grupo. Neste caso, tem-se três situações distintas: indivíduos com interação, sem interação ou com interação mediana ou transitória. A interação entre indivíduos do grupo é uma informação que reflete o quanto os membros interagem em determinado grupo, bem como

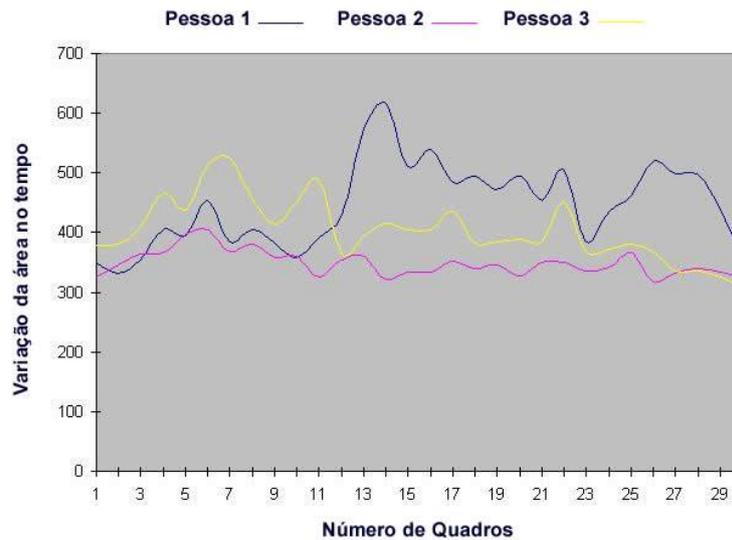


Figura 29: Gráfico mostrando a variação da área dos *blobs* no decorrer do tempo para três pessoas

uma forma de identificar qual indivíduo está interagindo com outro indivíduo do grupo. Uma interação entre duas pessoas é caracterizada quando as mesmas possuem seus vetores direção apontado diretamente uma para outra, valor de interação medido pelo produto interno gerado pelos vetores destas pessoas. Caso a métrica de interação I_{ij} seja próximo a 1, então é caracterizada interação entre os indivíduos. Por outro lado, se I_{ij} for próximo a -1 não existe interação. Por fim, se o valor for próximo a 0, a interação é caracterizada como mediana ou transitória.

Neste primeiro caso de estudo (A), a interação entre a pessoa 1 e 2 foi de 1,77%, com uma interação mediana de 77,88%. A interação entre a pessoa 2 e 3 foi de 1,10%, com interação mediana de 55,01%. Por fim, a interação entre pessoa 3 e 1 foi de 1,58%, com interação mediana de 62,20%. Pelo dados extraídos, é possível observar que em todo o tempo da cena em que o grupo ficou formado, existiu interação entre os indivíduos.

Com os dados extraídos do grupo, pode-se classificá-lo, levando em conta a maior porcentagem encontrada em cada item, segundo as seguintes características:

- grupo formado por três indivíduos
- geometria padrão: triângulo escaleno

- **classificação: grupo voluntário**
- **pessoas 1, 2 e 3 são ativas no grupo que compõem**
- **existe interação entre todos os membros do grupo**

Existem ainda outras variáveis que não foram citadas neste estudo de caso, que também são extraídas dos grupos e que poderiam servir de base para classificação de outras características, como localização/movimentação do grupo durante o vídeo (coordenadas (x,y)), entre outros, essas variáveis são as seguintes:

- centróide do grupo em x e y ;
- posição (x_i, y_i) da i -ésima da pessoa;

5.2 Caso B: Formação dinâmica de dois grupos

Neste segundo estudo de caso (B), existe a formação de dois grupos durante a amostragem de tempo. Tanto o grupo 1 como o grupo 2 são constituídos por duas pessoas, conforme mostra a Figura 30. Vale lembrar que os três estudos de caso apresentados neste trabalho foram realizados utilizando câmeras de topo. Além disso, foram filmados no mesmo ambiente, para que os dados a ser comparados entre os grupos possuíssem as mesmas características.

Os dados extraídos primeiramente para o grupo 1 foram os seguintes:

- tempo (duração do agrupamento): 110 quadros;
- número de pessoas que formam o grupo: duas pessoas = 100%;
- geometria formada: linha = 100%;
- classificação do grupo: voluntário = 100%;

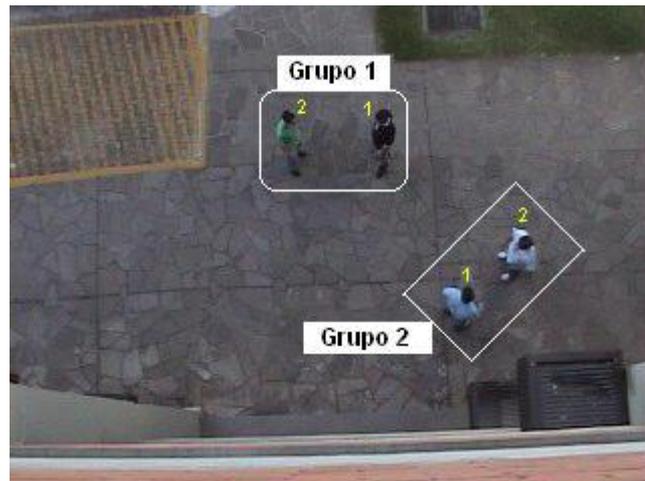


Figura 30: Identificação dos grupos detectados na cena no estudo de caso B

- *blob* da pessoa 1: *blobs* normais = 100%, *blobs* fundidos = 0%, variação média do *blob* = 27,81 pixels;
- *blob* da pessoa 2: *blobs* normais = 100%, *blobs* fundidos = 0%, variação média do *blob* = 14,86 pixels;
- interação entre pessoa 1 e 2: com interação = 100%;

Como este grupo foi composto somente por dois indivíduos do início ao final da cena, a porcentagem é de 100%, definido como um grupo formado por duas pessoas. Da mesma forma, por tratar-se sempre de dois pontos do início ao final do vídeo, a formação geométrica padrão detectada foi uma linha.

A classificação do grupo foi definida como grupo voluntário, devido ao grupo ter chegado a uma distância íntima quando já estava em estado estacionário e pelos vetores dos indivíduos que compõem o grupo estarem apontando um para o outro. Por permanecerem desta forma até o final da cena, constituem em 100% do tempo a classificação do grupo como voluntário.

O *blob* da pessoa 1, diferentemente do estudo de caso A, não teve fusão com nenhum outro *blob* durante o decorrer da cena, devido as pessoas que compõem este grupo não estarem em uma posição tão próxima como foi detectado no grupo do estudo de caso

A. Desta forma, o *blob* da pessoa 1 manteve-se normal durante todo o decorrer da cena, alcançando a taxa de 100%.

Quanto à variação média da área do *blob* da pessoa 1, existe algo interessante a ser considerado. A variação média encontrada para o *blob* da pessoa 1 foi de 27,81 pixels, onde é possível verificar no vídeo que esta pessoa faz alguns movimentos com o corpo durante o diálogo com a pessoa 2 que também compõe o grupo. No *blob* da pessoa 2, assim como ocorreu com o *blob* da pessoa 1, não houve *blobs* fundidos, pois a pessoa 2 manteve-se a uma distância não tão próxima a ponto de os dois *blobs* fundir-se, alcançado 100% de *blobs* normais durante todo o transcorrer da cena. No quesito variação da área do *blob* da pessoa 2, o valor encontrado foi de 14,86 pixels, valor baixo se comparado com os valores encontrados na variação dos *blobs* das pessoas 1, 2 e 3 que formam o grupo do estudo de caso A, onde os valores eram um pouco maiores que o dobro em relação ao encontrado na pessoa 2. De fato, estas informações fazem sentido se verificar-se a movimentação que teve a pessoa 2 no decorrer do vídeo. A pessoa 2 em questão manteve-se praticamente estática durante todo o decorrer da cena. Devido a esta postura, a variação da área do *blob* da pessoa 2 traz resultados coerentes.

Quanto aos níveis de interação entre a pessoa 1 e 2, a métrica de interação entre os indivíduos que compõem o grupo esteve sempre muito próximo a 1, desta forma a interação entre as pessoas do grupo manteve-se durante todo o decorrer da cena, totalizando 100% de interação entre a pessoa 1 e 2.

Com relação ao grupo 2 deste estudo de caso B, o grupo foi formado também por duas pessoas. Os dados extraídos pelo protótipo para o grupo 2, deste estudo de caso, foram os seguintes:

- tempo (duração do agrupamento): 121 quadros;
- número de pessoas que formam o grupo: duas pessoas = 100%;
- geometria formada: linha = 100%;

- classificação do grupo: voluntário = 100%;
- *blob* da pessoa 1: *blobs* normais = 100%, *blobs* fundidos = 0%, variação média do *blob* = 20,19 pixels;
- *blob* da pessoa 2: *blobs* normais = 100%, *blobs* fundidos = 0%, variação média do *blob* = 14,60 pixels;
- interação entre pessoa 1 e 2: com interação = 50,82%, interação mediana e/ou transitória = 49,18%;

Este trecho de vídeo para o grupo 2, teve duração de 121 quadros. Como não existiu um terceiro indivíduo neste grupo, o mesmo foi formado por duas pessoas durante todo o tempo. Por se tratar de um grupo de duas pessoas a geometria padrão linha manteve-se por toda a cena. O grupo 2, foi classificado como voluntário durante 100% do tempo da cena.

Com relação a variação média do *blob* da pessoa 1 e 2, os valores extraídos foram respectivamente de 20,19 e 14,6 *pixels*. Estes valores mais uma vez fazem sentido se comparados a outros grupos. As duas pessoas deste grupo mantem-se praticamente estáticas durante toda a cena. Apenas a movimentação de cabeça pode ser observada no vídeo por parte de um dos indivíduos.

A interação entre a pessoa 1 e 2 neste grupo foi de 50.82% (caracterizado como interação forte) e 49,18% (interação mediana ou transitória) durante a cena.

Com estes dados extraídos para os dois grupos do estudo de caso B, pode-se classificá-los da seguinte forma;

Grupo 1 - Estudo de caso B

- **grupo formado por duas pessoas**
- **geometria padrão: linha**

- **classificação: grupo voluntário**
- **pessoa 1 ativa e pessoa 2 não ativa no grupo**
- **interação existente entre a pessoa 1 e 2**

Grupo 2 - Estudo de caso B

- **grupo formado por duas pessoas**
- **geometria padrão: linha**
- **classificação: grupo voluntário**
- **pessoas 1 e 2 não são ativas no grupo**
- **interação existente entre a pessoa 1 e 2**

5.3 Caso C: Grupos compostos de duas e três pessoas em situações distintas

O estudo de caso C, apresenta dois grupos com formações distintas. O grupo 1, formado por três indivíduos e o grupo 2, formado por dois indivíduos, conforme mostra a Figura 31. Para o grupo 1, as informações capturadas e extraídas pelo protótipo foram as seguintes:

- **tempo (duração do agrupamento): 180 quadros;**
- **número de pessoas que formam o grupo: duas pessoas = 16,67%, três pessoas = 83,33%;**
- **geometria formada: linha = 16,67%, triângulo escaleno 83,33%;**
- **classificação do grupo: voluntário = 93,33%, involuntário 6,67%;**

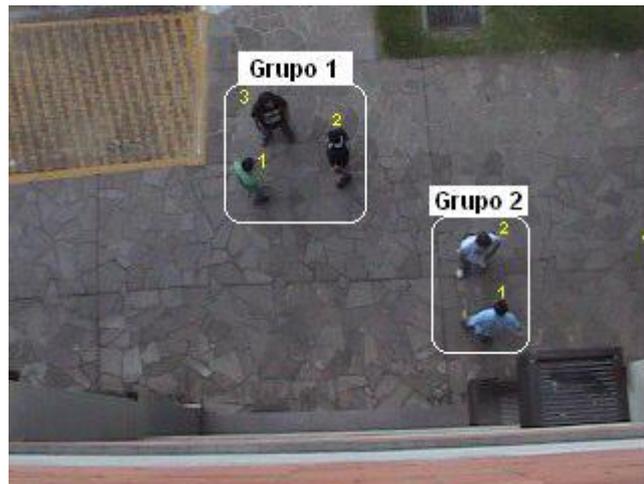


Figura 31: Identificação dos grupos no estudo de caso C

- *blob* da pessoa 1: *blobs* normais = 100%, variação média do *blob* = 12,74 pixels;
- *blob* da pessoa 2: *blobs* normais = 100%, variação média do *blob* = 21,17 pixels;
- *blob* da pessoa 3: *blobs* normais = 100%, variação média do *blob* = 13,39 pixels;
- interação entre pessoa 1 e 2: sem interação = 16,66%, interação mediana = 17,22%, com interação = 66,12%;
- interação entre pessoa 2 e 3: sem interação = 64,44%, interação mediana = 35,56%, com interação = 0%;
- interação entre pessoa 3 e 1: sem interação = 79,44%, interação mediana = 15%, com interação = 5,56%.

O tempo desta cena para o grupo 1, foi de 180 quadros. O grupo foi formado durante 16,67% do tempo por dois indivíduos, mais especificamente entre os indivíduos com ID 1 e 3, por estarem em uma distância íntima antes da pessoa 3 ser incorporada no grupo. Durante 83,33% da cena o grupo foi constituído por três pessoas.

Quanto à classificação do grupo, em 6,67% da cena o grupo formado foi classificado como involuntário, isto ocorreu devido ao grupo que estava se formando estar ainda em movimento, com um PPS pequeno. Desta forma a classificação de (JACQUES et al., 2007)

identificou o grupo como involuntário. Porém, em 93,33% da cena o grupo formado foi classificado como voluntário.

Neste grupo 1 do estudo de caso C, não houve fusão entre nenhum *blob*, sendo que a variação da área do *blob* para as pessoas 1, 2 e 3 respectivamente foi de 12.74, 21.17 e 13.39. De fato, é possível observar que a área do *blob* das pessoas 1 e 3 teve uma variação baixa devido a não movimentar-se durante a cena. A pessoa 3 teve uma participação um pouco maior, mesmo assim baixa se comparada aos valores encontrados em indivíduos que participam ativamente de outros grupos.

Quanto aos níveis de interação existiu uma interação maior entre os indivíduos com ID 1 e 2, principalmente por estarem de frente um para o outro. Considerando este fato, pode-se observar que a interação entre a pessoa 1 e 2 foi de 83,34%. Já a interação entre a pessoa 2 e 3 foi relativamente baixa, cerca de 35,56%. Por fim, a interação entre a pessoa 3 e 1 foi ainda menor 20,56%.

Com relação ao grupo 2 do estudo de caso C, este grupo foi formado por dois indivíduos, onde foram extraídas as seguintes informações pelo protótipo:

- tempo (duração do agrupamento): 153 quadros;
- número de pessoas que formam o grupo: duas pessoas = 100%;
- geometria formada: linha = 100%;
- classificação do grupo: voluntário = 88,13%, involuntário 11,87%;
- *blob* da pessoa 1: *blobs* normais = 100%, variação média do *blob* = 26,18 pixels;
- *blob* da pessoa 2: *blobs* normais = 100%, variação média do *blob* = 20,12 pixels;
- interação entre pessoa 1 e 2: com interação = 83,19%;

O tempo de duração deste grupo 2 foi de 153 quadros. Duas pessoas formaram o grupo durante todo o tempo da cena, com isso a geometria formada foi uma linha.

O grupo foi classificado como voluntário em 88,13% da cena, quando ficou em estado estacionário. Porém foi classificado como involuntário durante 11,87% da cena quando ainda estava em formação e em movimento. A interação entre a pessoa 1 e 2 foi constante (com interação) durante a maior parte da cena.

Unindo as informações do grupo 1 e 2 deste estudo de caso C, pode-se gerar a classificação geral destes grupos, conforme segue:

Grupo 1 - Estudo de caso C

- grupo formado por três pessoas
- geometria padrão: triângulo escaleno
- classificação: grupo voluntário
- pessoas 1, 2 e 3 não são ativas no grupo que compõem
- existe interação entre a pessoa 1 e 2 do grupo

Grupo 2 - Estudo de caso C

- grupo formado por duas pessoas
- geometria padrão: linha
- classificação: grupo voluntário
- pessoas 1 e 2 não são ativas no grupo que compõem
- interação existente entre pessoa 1 e 2

Capítulo 6

Resultados Complementares

Complementarmente ao assunto principal deste trabalho, foi realizado um estudo para avaliação do nível de atenção de um grupo com relação a um ponto de interesse proposto, como projeto do laboratório CROMOS. Este trabalho teve como principal objetivo realizar a classificação de grupo, extraindo informações como: distância, orientação e número de indivíduos no grupo. Através de um protótipo em Matlab, foi medido o nível de atenção de um grupo com relação a um ponto de interesse proposto.

A medida de atenção foi modelada a partir das seguintes variáveis:

- a orientação de um grupo com relação a um ponto de interesse proposto;
- a distância de um grupo na cena com relação a um ponto de interesse proposto definido previamente;
- o número de pessoas presentes na cena para detecção de um grupo;
- o tempo decorrente da cena, para extrair informações de atenção de grupo com relação ao ponto de interesse exposto;
- a geometria formada pelo grupo, bem como suas possíveis subdivisões;

Para avaliar o nível de atenção, foram utilizadas duas seqüências de vídeo: a



Figura 32: Ambiente de realização dos vídeos



Figura 33: Isolando variáveis no modelo de atenção

primeira, filmada ao lado da Biblioteca Central da UNISINOS; a segunda, em frente ao prédio do PIPCA (Programa Interdisciplinar de Pós-Graduação em Computação Aplicada). Ambas filmagens compartilharam de características comuns:

- câmera oblíqua (localizada em cima do ponto de interesse, neste caso tratando-se de um *banner* exposto).
- pessoas localizadas em três distâncias pré-definidas com relação ao objeto de interesse (ver Figura 32).
- número variável de pessoas olhando simultaneamente para o ponto de atenção (ver Figura 33).

Para que as variáveis do modelo de atenção pudessem ser medidas de forma a calcular da melhor maneira possível a atenção do grupo com relação a um ponto de interesse, foi criado um questionário reunindo a opinião de diversas pessoas sobre a importância de cada uma destas variáveis no modelo. O questionário visou avaliar a importância das

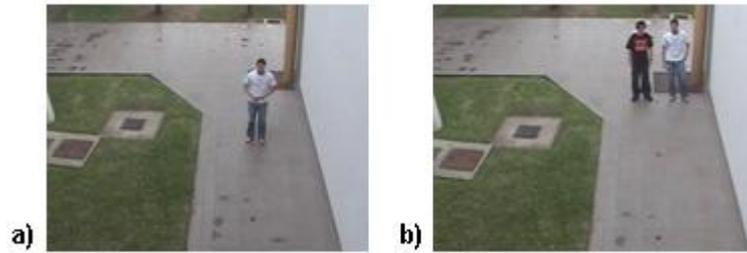


Figura 34: Comparação entre as variáveis distância e número de pessoas

variáveis **distância**, **orientação** e **número de pessoas** de forma individual e combinada. A pergunta a ser respondida em todas as questões foi a seguinte: *Ordene as três imagens abaixo, com relação à atenção dada ao seu produto*, e foram mostradas para os sujeitos várias combinações com número de pessoas variável, distâncias variáveis e orientação com relação ao produto variável. O sujeito deveria ordenar as imagens da maior atenção à menor. Um exemplo ilustrando duas dessas imagens é mostrado na Figura 34.

Tendo como base as respostas das pessoas (26 no total) foi possível estimar-se uma ordenação quanto à importância das variáveis:

- 1º - Orientação;
- 2º - Distância;
- 3º - Número de Pessoas.

As respostas dos sujeitos que utilizaram o sistema apresentou um consenso em relação a esta ordem. Na verdade, pesquisas mais avançadas são necessárias a fim de estimar estes números com maior realismo.

Também foi criada uma interface gráfica que mostra para o usuário o nível de atenção com relação à orientação, distância e número de pessoas no grupo. A Figura 35 ilustra um exemplo da interface, onde se pode observar informações do sistema, tais como:

- tempo decorrido da cena (canto superior esquerdo);
- a distância do grupo com relação ao ponto de interesse (canto superior esquerdo);



Figura 35: Imagem da execução do modelo e extração de informações do grupo

Informações relativas a orientação, distância e número de pessoas, são extraídas da seguinte forma:

- a orientação do grupo na cena é extraída através de uma limiarização dos ângulos com valores entre 0 e 1;
- a distância do grupo com relação ao ponto de interesse é realizada através da verificação a cada quadro das coordenadas x e y de cada indivíduo, extraíndo assim a distância média do grupo com relação a coordenada (x,y) do ponto de interesse.
- o número de pessoas na cena varia de 1 à 3 considerando um valor máximo de pessoas fixado em três componentes;

Após o processamento destes arquivos e o final de suas execuções, o sistema pode gerar gráficos que mostram a evolução temporal dos parâmetros calculados. Por exemplo, a Figura 36 mostra a distância do grupo com relação ao objeto de atenção ao longo do tempo. Tais gráficos poderiam ser utilizados para avaliar, de acordo com diferentes parâmetros (orientação, distancia e número de pessoas), o impacto causado pelo produto em grupos de pessoas ao longo do tempo.

Deve-se levar em consideração que este capítulo mostra resultados complementares, com relação à medição do nível de atenção, sendo que vários pontos pode ser melhorados.

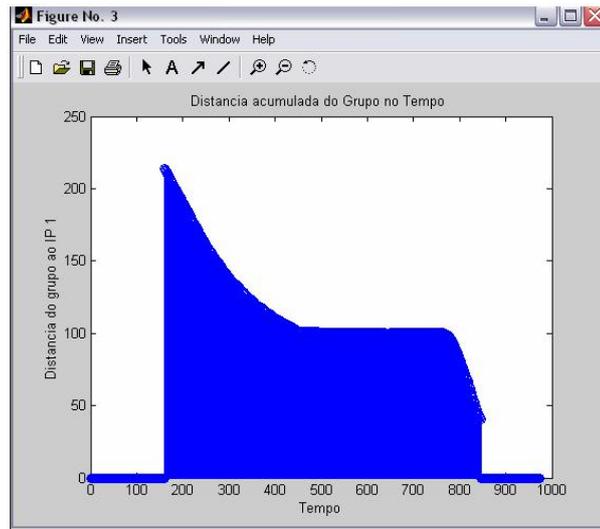


Figura 36: Gráfico de distância do grupo em relação ao objeto de interesse ao longo do tempo.

Por exemplo, a orientação de cada pessoa é obtida pelo acompanhamento, o que pode gerar problemas para pessoas estáticas ou com pouco movimento. Além disso, a estimativa do vetor de orientação é em duas dimensões e em coordenadas de imagem, quando o ideal seria 3D e em coordenadas de mundo. Para isso, seria necessária uma calibração da câmera empregada nos testes.

Capítulo 7

Conclusão

Este trabalho apresentou um modelo de classificação de grupos utilizando informações de geometria e detecção de atividades intragrupo, com o objetivo específico de poder extrair padrões de classificação de grupos de indivíduos em seqüências de vídeo. Como embasamento teórico, foram abordados nesta pesquisa três tópicos fundamentais, que envolveram em sua ordem: sistemas computacionais de detecção e acompanhamento, a análise do comportamento humano baseado nas interações interpessoais e a análise de aplicações envolvendo grupos de robôs.

O modelo proposto desenvolvido neste trabalho foi aplicado em um protótipo para detectar e classificar grupos presentes em uma seqüência de vídeo. Poderá ser utilizado também para analisar o comportamento das pessoas, as distâncias que mantêm umas das outras, suas orientações quando distribuídas em grupos, a geometria que formam, além de informações quanto à movimentação dos indivíduos dentro de um grupo, medidas de atividade dos membros dos grupos, dentre outras informações que possam ser relevantes à aplicação.

Com relação aos estudos de caso analisados neste trabalho, os três experimentos foram realizados utilizando-se de câmera de topo na mesma posição e distância na filmagem dos três vídeos. Assim, os dados gerados nesses vídeos puderam ser comparados

entre si, sem a necessidade de ajustes de limiares.

No primeiro experimento (caso A), existia um grupo formado por três pessoas, todas interagindo entre si. A grande expectativa neste primeiro experimento era saber a variação da área do *blob* das pessoas durante o decorrer da cena, para comparar posteriormente com grupos onde as pessoas não estavam interagindo entre si, ou onde as pessoas do grupo estavam visualmente paradas (sem movimentar-se).

No segundo experimento (caso B) tinha-se dois grupos formados por duas pessoas em cada grupo. Nestes grupos era possível observar que as pessoas movimentavam-se menos que no primeiro experimento (caso A). De fato, o protótipo detectou que no primeiro grupo existia uma pessoa classificada como ativa no grupo (gesticulando, movimentando-se) e uma pessoa não ativa no grupo (sem movimento ou parada). No grupo 2, porém, a classificação final encontrada foi que as duas pessoas não estavam ativas no grupo, como de fato é possível observar no vídeo que a movimentação das pessoas é praticamente nula.

O estudo de caso C visou reunir dois grupos com formações distintas, com o número máximo de pessoas para formação de grupo, conforme proposto neste trabalho. O número de pessoas, movimentação e interação entre os indivíduos do grupo foram variados para que se pudesse chegar a resultados comparativos entre os estudos de caso A, B e C.

Analisando os resultados gerados pelo protótipo de maneira geral para os estudos de caso A, B e C, chegou-se algumas conclusões interessantes, como:

- a variação da área do *blob* aparenta ser um indicador coerente com relação a atividade intragrupo, assim podendo ser utilizada para detectar os grupos com maior movimentação (ou gesticulando) entre seus indivíduos;
- a interação entre os indivíduos do grupo, mesmo que realizada algumas vezes de forma semi-automática, traz informações coerentes quando busca-se saber exatamente quais são as pessoas que estão interagindo durante determinado tempo no

decorrer da cena;

- o fato das informações serem armazenadas quadro-a-quadro, chegando a um total de vinte e três variáveis diferentes, geram uma grande quantidade de dados que podem ser utilizados para uma análise mais detalhada e posterior classificação do grupo.

Salienta-se que este trabalho focou em grupos de duas ou três pessoas por estes ocorrerem com mais frequência no dia-a-dia. De fato, segundo um estudo realizado por (SOMMER, 1973), verificou-se que 71% dos grupos (tanto informais quanto de trabalho) continham apenas dois indivíduos, 21% continham três indivíduos, 6% continham quatro indivíduos e apenas 2% continham cinco ou mais indivíduos.

Deve-se considerar ainda, que muitas das variáveis do sistema foram pouco exploradas na classificação dos grupos e medida de atividade, e que podem ser utilizadas como trabalhos futuros. Por exemplo, a geometria do grupo formada aliada à relações de interações entre os membros pode melhorar a detecção de atividade intragrupo.

Referências

- BARFOOT, T. D.; CLARK, C. M. Motion planning for formations of mobile robots. In: *Robotics and Autonomous Systems*. Stanford, USA: Elsevier Science Inc., 2003. p. 65–78.
- BELEZNAI, C.; FRÜHSTÜCK, B.; BISCHOF, H. Human detection in groups using a fast mean shift procedure. In: *ICIP 04 - Image Processing*. Singapore: IEEE International Conference On Image Processing, 2004. p. 349–352. ISBN 0-7803-8554-3.
- COLOMBARI, A.; FUSIELLO, A.; MURINO, V. Segmentation and tracking of multiple video objects. *Pattern Recognition*, Elsevier Science Inc., New York, NY, USA, v. 40, n. 4, p. 1307–1317, 2007. ISSN 0031-3203.
- CUI, J. et al. Laser-based detection and tracking of multiple people in crowds. *Computer Vision Image Understanding*, Elsevier Science Inc., New York, NY, USA, v. 106, n. 2-3, p. 300–312, 2007. ISSN 1077-3142.
- DU, Y. et al. Recognizing interaction activities using dynamic bayesian network. In: *ICPR '06: Proceedings of the 18th International Conference on Pattern Recognition*. Washington, DC, USA: IEEE Computer Society, 2006. p. 618–621. ISBN 0-7695-2521-0.
- FUENTES, L.; VELASTIN, S. People tracking in surveillance applications. *Image and Vision Computing*, Elsevier Science Inc., New Zealand, v. 24, n. 11, p. 1165–1171, 2005. ISSN 0262-8856.
- HALL, T. E. *A dimensão oculta*. Rio de Janeiro: Tradução: Sônia Coutinho do original em inglês: The hidden dimension (1989), 1989. 136 p.
- HOSIE, R.; VENKATESH, S.; WEST, G. Classifying and detecting group behaviour from visual surveillance data. In: *ICPR '98: Proceedings of the 14th International Conference on Pattern Recognition-Volume 1*. Washington, DC, USA: IEEE Computer Society, 1998. p. 602. ISBN 0-8186-8512-3.
- IKEMOTO, Y. et al. Gradual spatial pattern formation of homogeneous robot group. *Inf. Comput. Sci.*, Elsevier Science Inc., New York, NY, USA, v. 171, n. 4, p. 431–445, 2005. ISSN 0020-0255.
- JACQUES, C. S. J. J. *Utilizando Visão Computacional para Simular e Validar Comportamentos de Multidões de Humanos Virtuais*. São Leopoldo - RS, Brasil: Dissertação (Mestrado) - Universidade do Vale do Rio dos Sinos (UNISINOS), 2006. 110 p.

- JACQUES, C. S. J. J. et al. Understanding people motion in video sequences using voronoi diagrams: Detecting and classifying groups. *Pattern Analysis and Applications*, Springer-Verlag, London, UK, v. 10, n. 4, p. 321–332, 2007. ISSN 1433-7541.
- KANG, H.; KIM, D.; BANG, S. Y. Real-time multiple people tracking using competitive. In: *ICPR '02: Proceedings of the 16 th International Conference on Pattern Recognition (ICPR'02) Volume 1*. Washington, DC, USA: IEEE Computer Society, 2002. p. 10413. ISBN 0-7695-1695-X.
- KEPHART, W. M. *Quantitative Analysis of Intra-Group Relationships*. New York, USA: American J. Sociology, 1950. 165 p.
- KRISHNANAND, K. N.; GHOSE, D. Formations of minimalist mobile robots using local-templates and spatially distributed interactions. *Robotics and Autonomous Systems*, Elsevier Science Inc., Bangalore, India, v. 53, n. 3-4, p. 194–213, 2005.
- MARCENARO, L.; MARCHESOTTI, L.; REGAZZONI, C. S. Self-organizing shape description for tracking and classifying multiple interacting objects. *Computer Vision, Graphics and Image Processing*, 5th Indian Conference, Icvqip, Madurai, India, v. 24, n. 11, p. 1179–1191, 2006.
- MATARIC, J. M. Designing and understanding adaptive group behavior. *International Society of Adaptive Behavior*, Computer Science Department, v. 4, n. 1, p. 51–80, 1995.
- MATARIC, M. Issues and approaches in the design of collective autonomous agents. *Robotics and Autonomous Systems*, Elsevier Science Inc., v. 16, p. 321–331, 1994.
- MCKENNA, J. S. et al. Tracking groups of people. *Computer Vision and Image Understanding: CVIU*, Elsevier Science Inc., Kingston, UK, v. 80, n. 1, p. 42–56, 2000.
- MILLS, M. T. *Sociologia dos pequenos grupos*. Universidade de Yale: Tradução de Dante Moreira Leite do original: *The Sociology of Small Groups*, 1970. 179 p.
- PAYTON, D. W.; ESTKOWSKI, R.; HOWARD, M. Compound behaviors in pheromone robotics. *Robotics and Autonomous Systems*, Elsevier Science Inc., v. 44, n. 3-4, p. 229–240, 2003.
- REIF, J. H.; WANG, H. Social potential fields: a distributed behavioral control for autonomous robots. In: *WAFR: Proceedings of the workshop on Algorithmic foundations of robotics*. Natick, MA, USA: A. K. Peters, Ltd., 1995. p. 331–345. ISBN 1-56881-045-8.
- SILVA, M. Interatividade: uma mudança fundamental do esquema clássico da comunicação. In: *Intercom - Sociedade Brasileira de Estudos Interdisciplinares da Comunicação*. Santos - SP, Brasil: XXX Congresso Brasileiro de Ciências da Comunicação., 2007. p. 125–138.
- SOMMER, R. *Espaço Pessoal*. São Paulo - SP, Brasil: Tradução: Dante Moreira Leite do original: *Personal Space: the behavioral basis of design.*, 1973. 92 p.
- WANG, H.; D., S. A consensus-based method for tracking: Modelling background scenario and foreground appearance. *Pattern Recognition*, Elsevier Science Inc., New York, NY, USA, v. 40, n. 3, p. 1091–1105, 2007. ISSN 0031-3203.

WANG, L.; HU, W.; TAN, T. Recent developments in human motion analysis. In: *Pattern Recognition*. [S.l.]: Elsevier Science Inc., 2003. p. 585–601.

YAMAGUCHI, H.; ARAI, T.; BENI, G. A distributed control scheme for multiple robotic vehicles to make group formations. *Robotics and Autonomous Systems*, Elsevier Science Inc., v. 36, n. 4, p. 125–147, 2001.

YILMAZ, A.; JAVED, O.; SHAH, M. Object tracking: A survey. *ACM Computer Survey*, ACM Press, New York, NY, USA, v. 38, n. 4, p. 13, 2006. ISSN 0360-0300.