

## 4 Avaliação Experimental

Um protótipo que implementa o modelo proposto foi construído para avaliar o desempenho do sistema, utilizando a linguagem MatLab. Diferentes seqüências de imagens 2D, de 160 x 320 pixels, obtidas através de uma webcam convencional modelo PD 1001 da Creative, foram utilizadas nesta avaliação. Algumas destas seqüências contém situações simuladas que foram planejadas de modo a permitir a avaliação de cada passo do algoritmo isoladamente. Já outras seqüências foram gravadas em ambientes onde não havia controle algum sobre as pessoas ou sobre a iluminação, ou seja, situações ilustrativas de condições reais de aplicação do sistema.

Este capítulo apresenta e descreve o plano experimental elaborado e discute os resultados obtidos. Está dividido em quatro seções. A primeira descreve todos os ambientes envolvidos na avaliação. A segunda descreve as situações simuladas assim como os resultados alcançados em cada uma delas. Já a terceira apresenta os resultados obtidos nas situações reais, onde não houve controle de variável alguma, e os relaciona com os objetivos propostos por este trabalho. Por fim, a quarta e última seção avalia a sensibilidade dos parâmetros envolvidos no funcionamento do sistema, assim como esclarece o significado de cada parâmetro e apresenta seus valores.

### 4.1. Descrição dos Ambientes

Foram utilizados quatro ambientes distintos para as filmagens das situações simuladas e um ambiente para a filmagem da situação real. A

Tabela 1 apresenta algumas características de cada um destes ambientes, no que se refere ao tipo de ambiente (coberto ou ao ar livre), ao tipo de iluminação (natural ou artificial), às condições do ambiente (ordenado ou desordenado) e ao tipo de filmagem realizado (simulada ou real).

## 4.2. Experimentos com Seqüências de Vídeo de Situações Simuladas

Esta primeira subseção se divide nas etapas do processo: separação do fundo / primeiro plano, atualização do fundo, localização / contagem de pessoas e detecção de atitudes suspeitas. Para cada uma destas fases é apresentado o procedimento desenvolvido para a avaliação do sistema proposto.

Tabela 1: Descrição dos ambientes.

		<b>Ambiente</b>	<b>Iluminação</b>	<b>Condições</b>	<b>Filmagem</b>
<b>Ambiente 1</b>		Coberto.	Natural.	Desordenado.	Simulada.
<b>Ambiente 2</b>		Ao ar livre.	Natural.	Ordenado.	Simulada.
<b>Ambiente 3</b>		Ao ar livre.	Natural.	Um pouco desordenado.	Simulada.
<b>Ambiente 4</b>		Coberto.	Natural.	Ordenado.	Simulada.
<b>Ambiente 5</b>		Coberto.	Natural e artificial.	Ordenado, porém sobrecarregado.	Real.

## **4.2.1. Separação do Fundo / Primeiro Plano**

Nesta etapa avalia-se o comportamento do sistema com relação a sombras. A seguir é apresentado o procedimento adotado para a avaliação experimental deste tópico, assim como os correspondentes resultados.

### **4.2.1.1. Sombras**

Para a avaliação de sombras foram elaboradas 6 seqüências de 120 quadros cada (30 segundos), 3 no ambiente 1 e 3 no ambiente 2. Em cada uma delas, a situação era igual: o ambiente iniciava sem ninguém (Figura 12a e Figura 12b), em seguida uma pessoa entrava com um objeto (Figura 12c e Figura 12d), o colocava em uma determinada posição do ambiente (Figura 12e e Figura 12f) e saía (Figura 12g e Figura 12h).

Na medida em que o objeto era deixado em uma determinada posição, ele projetava uma sombra no chão, que, de acordo com a iluminação ambiente, podia ser suave, média ou forte. Para que isto fosse possível, em cada ambiente, foram utilizados 3 níveis de iluminação distintos conforme mostrado na Figura 13. O primeiro deles fez uso apenas da iluminação ambiente e foi responsável pela obtenção da sombra suave (Figura 13a e Figura 13d). Já o segundo nível de iluminação foi gerado com auxílio de uma lâmpada de 150 watts e uma de 200 watts acesas em um dos lados do objeto para que do outro lado fosse projetada uma sombra média (Figura 13b e Figura 13e). Por fim, o terceiro nível de iluminação, foi obtido igual ao segundo, porém com uma lâmpada de 150 watts e duas de 200 watts, desta forma obteve-se a sombra forte (Figura 13c e Figura 13f).

Como pode ser observado na Figura 13, o objeto utilizado era propositalmente de uma cor única que contrastava com o fundo (cor branca). Desta forma, foi possível executar o seguinte procedimento.

- i) Pelo método proposto encontrou-se a região da imagem correspondente ao primeiro plano.
- ii) Desta região retirou-se os pixels de cor branca que, por sua vez, correspondiam ao objeto inserido no ambiente.
- iii) Após esta retirada, os pixels que sobraram correspondiam aos pixels que eram, de fato, sombra. Tais pixels foram contabilizados.

- iv) Em seguida, o método proposto foi aplicado na região encontrada como primeiro plano e obteve-se, então, a região que o sistema detectou como sombra. Os pixels de tal região também foram contabilizados.
- v) Por fim, foi calculada a razão entre o número de pixels contabilizado no passo iv) (sombra detectada pelo sistema) e o número de pixels contabilizado no passo iii) (sombra real). Tal razão corresponde ao percentual de sombra detectado pelo sistema proposto, ou seja, o percentual de acerto.

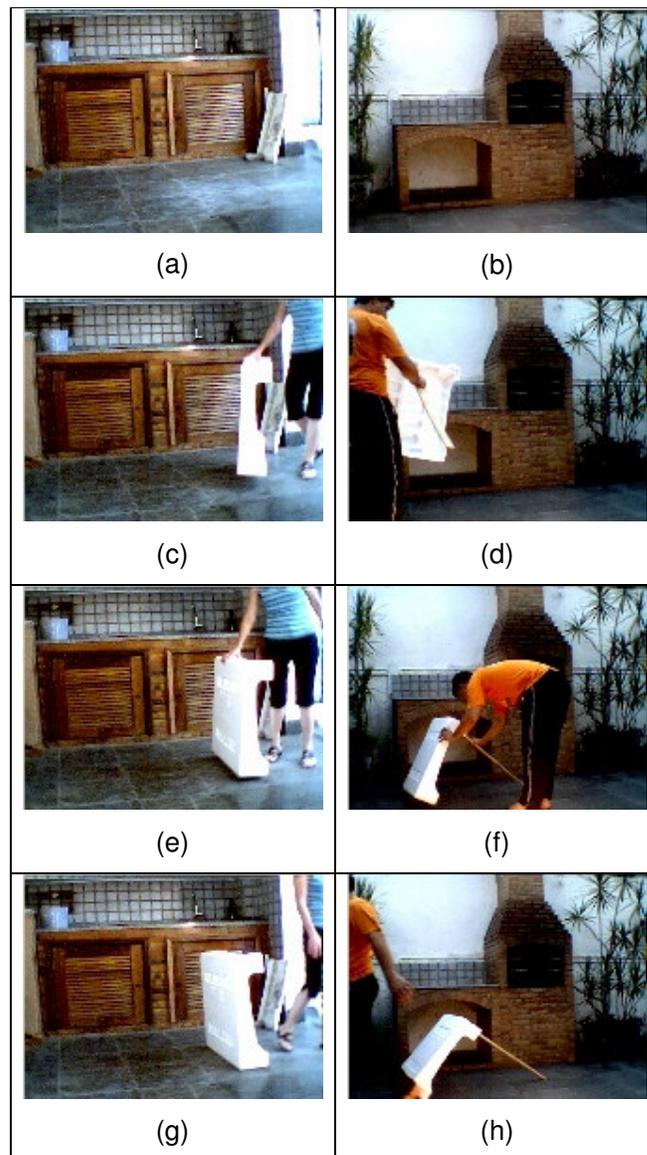


Figura 12: Sequências para análise de sombras: (a) ambiente 1 sem ninguém; (b) ambiente 2 sem ninguém; (c) e (d) pessoa entra com objeto; (e) e (f) pessoa deixa objeto no local; (g) e (h) pessoa sai.

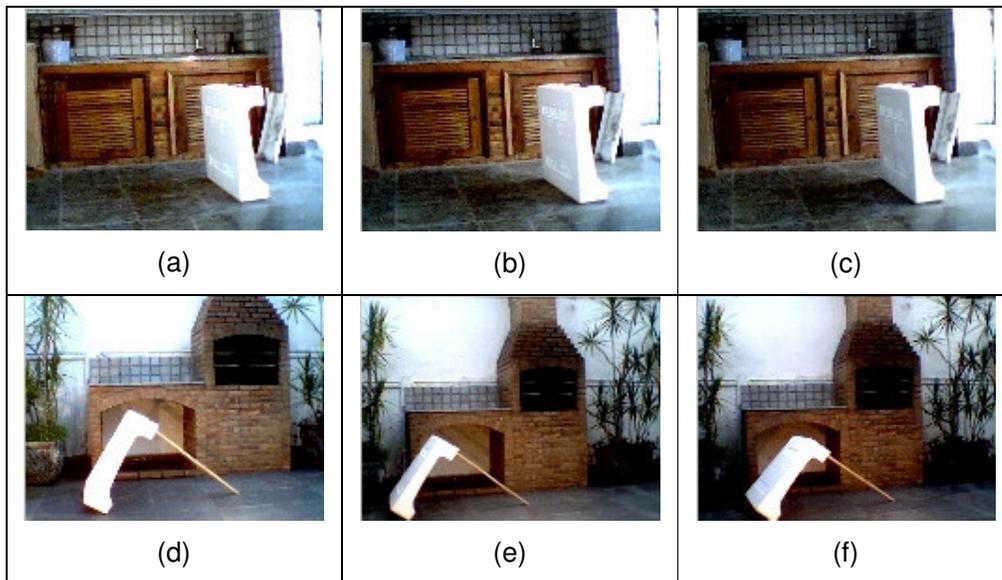


Figura 13: Sombras geradas: ambiente 1 – (a) suave; (b) média; (c) forte; ambiente 2 – (d) suave; (e) média; (f) forte.

Na Tabela 2 é indicado o percentual de acerto para cada uma das seqüências elaboradas, ou seja, o desempenho do método.

Tabela 2: Percentual de sombra detectado.

	<b>Ambiente 1</b>	<b>Ambiente 2</b>
<b><i>Suave</i></b>	82,07 %	83,95 %
<b><i>Média</i></b>	69,41 %	54,62 %
<b><i>Forte</i></b>	72,04 %	58,40 %

Para as sombras suaves, os resultados foram claramente melhores, uma vez que o algoritmo de detecção se baseia no contraste dos pixels em relação ao fundo e para este tipo de sombra o contraste é menor. Com isto, um percentual significativo dos pixels relativos a sombras é detectado. Já para sombras média e forte, o percentual de acerto não foi tão alto, já que o contraste destes tipos de sombra em relação ao fundo é maior. Além disto, o percentual detectado de sombra forte foi levemente maior que o de sombra média, quando o esperado era o contrário. Na verdade, fora o ajuste automático de ganho da câmera, a variação de iluminação de uma situação para a outra (diferença de uma lâmpada de 200 watts apenas) não foi muito significativa para o método, de forma que os resultados foram muito próximos.

De qualquer forma, os resultados condizem com o desejado, pois o comportamento desejado é que pixels muito diferentes do fundo sejam dados como primeiro plano, portanto, para sombras mais fortes, necessariamente, o acerto deve ser menor, caso contrário o primeiro plano não seria detectado de forma correta.

#### **4.2.2. Atualização da Estimativa do Fundo**

Neste experimento avalia-se a capacidade do método de absorver variações suaves no fundo, devidas, por exemplo, a alterações na iluminação, e variações abruptas, decorrentes de alterações permanentes, como por exemplo, a retirada ou colocação de objetos no fundo. Esta funcionalidade deve estar ativa inclusive na presença de pessoas em movimento na cena.

A seguir estão apresentados os resultados encontrados em cada uma destas situações.

##### **4.2.2.1. Presença de Pessoas**

Uma característica desejável do sistema é que a atualização de fundo seja feita sem prejuízo, independente da presença de pessoas circulando no ambiente. Caso contrário, o sistema não poderá ser atualizado continuamente, já que será necessário que não haja ninguém em cena para que a atualização seja efetuada, o que não é o esperado, pois, a princípio, as imagens são capturadas em ambientes de livre circulação.

Desta forma, o método proposto foi desenvolvido para atualizar o fundo independentemente da presença de pessoas no ambiente. O algoritmo elaborado só atualiza os pixels que são classificados como fundo ou que ficam estáveis no primeiro plano por mais de um determinado número de quadros. Portanto, basta que as pessoas não fiquem paradas por muito tempo no mesmo local para não interferirem na atualização dinâmica do fundo.

Três seqüências foram elaboradas para os testes de atualização de fundo na presença de pessoas. A Tabela 3 mostra por quantos quadros tais seqüências apresentaram uma ou nenhuma pessoa. Cada uma delas teve a duração de 240 quadros (1 minuto).

Tabela 3: Relação entre o número de pessoas e a quantidade de quadros das seqüências de teste de atualização do fundo nos casos de presença de pessoas.

Número de Pessoas	Quantidade de Quadros		
	<i>Ambiente 1</i>	<i>Ambiente 2</i>	<i>Ambiente 3</i>
0	27	23	45
1	213	216	195

Na Figura 14, pode-se observar as etapas de uma das seqüências elaboradas. Tais imagens são relativas à seqüência do ambiente 3. As regiões brancas das imagens da segunda coluna representam os primeiros planos relativos às respectivas imagens da primeira coluna.

Inicialmente, o ambiente começa sem a presença de qualquer pessoa (Figura 14a). Na Figura 14b pode-se observar que o primeiro plano está todo preto, ou seja, nenhuma região foi dada como primeiro plano. Já na Figura 14c uma pessoa entra em cena e na Figura 14d é possível ver que já se tem uma região relativa à esta pessoa no primeiro plano encontrado. Na Figura 14e a pessoa circula pelo ambiente e novamente é classificada como pertencente ao primeiro plano (Figura 14f). Por fim, a pessoa sai do ambiente (Figura 14g), mas a parte do corpo dela que é visível na imagem, ainda é classificada como primeiro plano (Figura 14h). Pelas imagens dos primeiros planos apresentados na Figura 14 pode-se observar que nestes exemplos as pessoas foram, realmente, rotuladas como parte do primeiro plano, estando, assim, de acordo com o desenvolvimento do método e não interferindo na atualização do fundo.

#### 4.2.2.2. Variações de Iluminação

Para avaliar a atualização do fundo frente a variações de iluminação, foram elaboradas duas condições de iluminação. A primeira onde o ambiente em análise se encontra mais claro (iluminação máxima) e a segunda onde se encontra mais escuro (iluminação mínima). A condição de iluminação máxima foi obtida através da utilização de 3 lâmpadas, uma de 150 watts e duas de 200 watts, além da iluminação natural. Já a condição de iluminação mínima foi obtida através da utilização de apenas uma lâmpada de 200 watts, além da iluminação natural. A Figura 15a, a Figura 15b e a Figura 15c apresentam a iluminação máxima dos ambientes 1, 2 e 3, respectivamente. Já a Figura 15d, a Figura 15e e a Figura 15f apresentam a iluminação mínima destes ambientes.

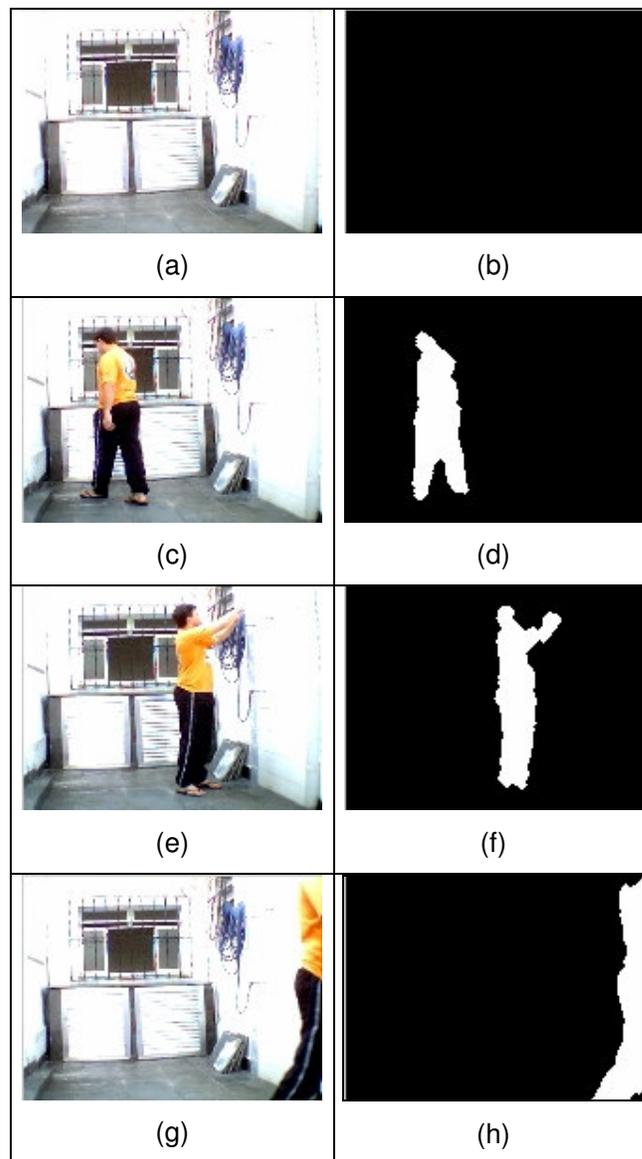


Figura 14: Seqüência para análise da atualização do fundo nos casos de presença de pessoas: (a) ambiente 3 sem ninguém; (b) primeiro plano correspondente; (c) pessoa entra; (d) primeiro plano correspondente; (e) pessoa circula; (f) primeiro plano correspondente; (g) pessoa sai; (h) primeiro plano correspondente.

Em seguida, foram elaboradas três situações distintas, onde tais condições de iluminação variavam, para serem implementadas nos ambientes 1, 2 e 3, conforme descrito a seguir.

- i) *Varição de iluminação suave*: durante 15 segundos (60 quadros) manteve-se a iluminação mínima e, em seguida, a iluminação máxima por 45 segundos (120 quadros), totalizando 1 minuto (240 quadros) de filmagem.

- ii) *Variação de iluminação média*: durante 15 segundos (60 quadros) manteve-se a iluminação mínima, em seguida, a iluminação máxima por 15 segundos (60 quadros), e, por fim, a mínima novamente por 30 segundos, totalizando 1 minuto (240 quadros) de filmagem.
- iii) *Variação de iluminação abrupta*: durante 15 segundos (60 quadros) manteve-se a iluminação mínima, em seguida, a iluminação máxima por 15 segundos (60 quadros), a mínima novamente por 15 segundos e, por fim, a máxima por mais 30 segundos, totalizando 1 minuto e 15 segundos (300 quadros) de filmagem.

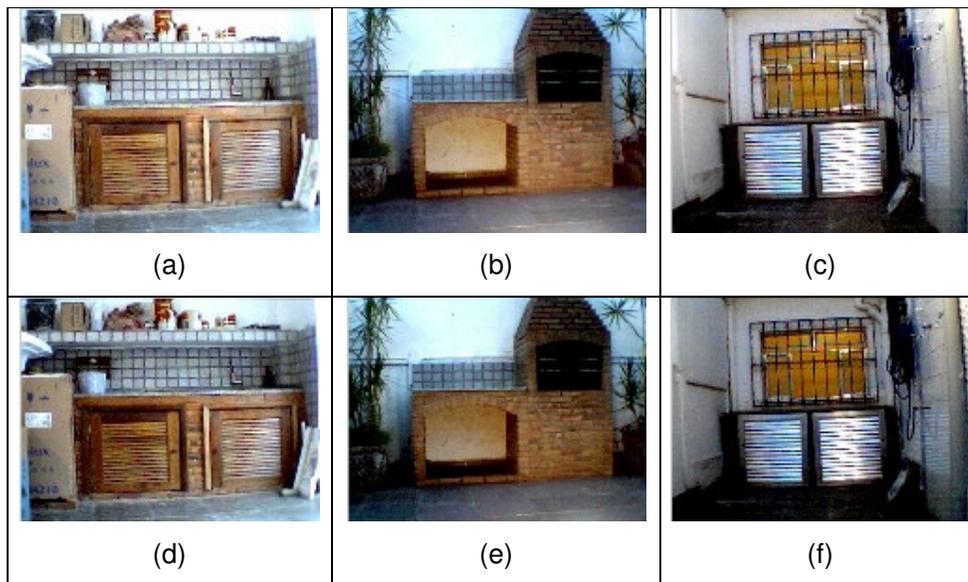


Figura 15: Condições de iluminação geradas: máxima – ambientes 1 (a), 2 (b) e 3 (c); mínima – ambientes 1 (d), 2 (e) e 3 (f).

Nos nove experimentos (3 variações para cada um dos 3 ambientes distintos), o procedimento de atualização do fundo acompanhou as variações realizadas perfeitamente. Isto é, as variações de iluminação foram sendo gradativamente incorporadas na atualização do fundo.

Na Figura 16 pode ser observado um exemplo da atualização de fundo relativa a uma variação da iluminação mínima à máxima no ambiente 2. A Figura 16a, a Figura 16b e a Figura 16c mostram uma seqüência de 3 quadros desta variação, onde na Figura 16a a iluminação é a mínima, ao passo que na Figura 16b e na Figura 16c a iluminação já é a máxima. Já a seqüência composta pela Figura 16d, Figura 16e e Figura 16f mostra as atualizações de fundo relativas, respectivamente, ao quadros anteriormente citados. Na Figura 16d a iluminação ainda é a mínima, enquanto na Figura 16e a atualização do

fundo passa por um estágio intermediário onde a iluminação se encontra entre a mínima e a máxima. Já na Figura 16f a atualização do fundo já atingiu a iluminação máxima desejada. Estas variações podem ser melhor observadas na região abaixo da bancada da churrasqueira.

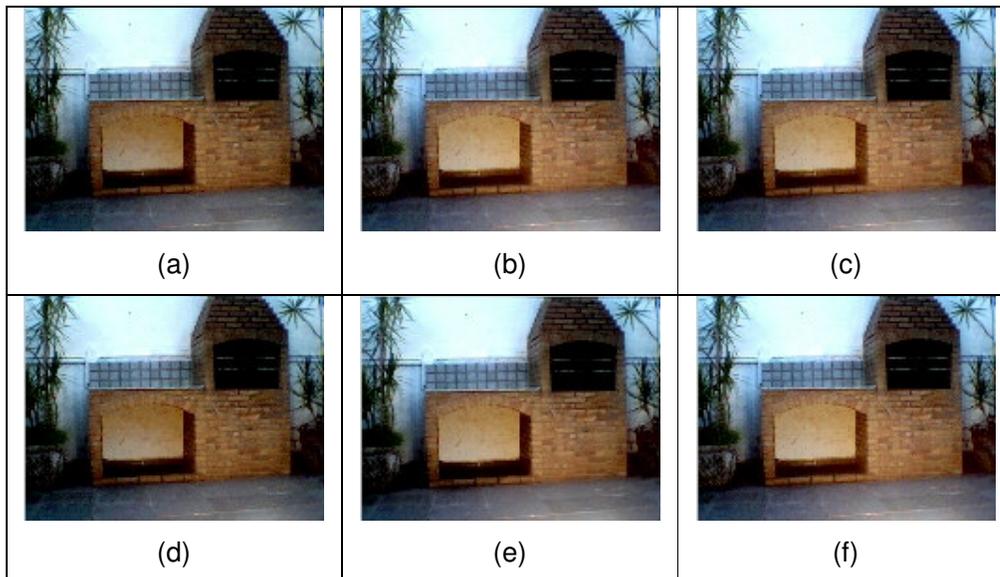


Figura 16: Variação de iluminação gerada: (a) mínima; (b) máxima; (c) máxima; e respectiva variação das atualizações de fundo: (a) mínima; (b) mediana; (c) máxima.

#### 4.2.2.3. Alterações Permanentes

Por último, foram feitas filmagens onde alterações permanentes foram introduzidas no fundo. Na Figura 17 está apresentado o primeiro caso de alteração permanente realizado no ambiente 4. Inicialmente, o ambiente se encontra sem a presença de ninguém (Figura 17a). Em seguida, uma pessoa entra na cena e leva (Figura 17b) um objeto, ocasionando uma alteração permanente no local. Por fim, a pessoa se retira do local (Figura 17c).

A Figura 18 apresenta as etapas da atualização da estimativa de fundo relativas a alteração permanente ilustrada na Figura 17. No momento em que a lixeira é retirada do local em que estava, o sistema passa a considerar este local como parte do primeiro plano, uma vez que tal região passa a diferir do fundo. Entretanto, temporariamente, a lixeira continua presente na estimativa de fundo (Figura 18a). Na medida em que o tempo em que a lixeira não se encontra mais no local em que estava, coincide com o tempo de atualização do fundo ( $K_{máx}$  quadros), o fundo é atualizado neste local. Observe que esta atualização pode

ocorrer de forma gradativa, em grupos de pixels (Figura 18b, Figura 18c, Figura 18d e Figura 18e), e não, necessariamente, de uma só vez. De uma forma ou de outra, ao final da atualização, a lixeira deixa de estar presente na estimativa do fundo (Figura 18f).

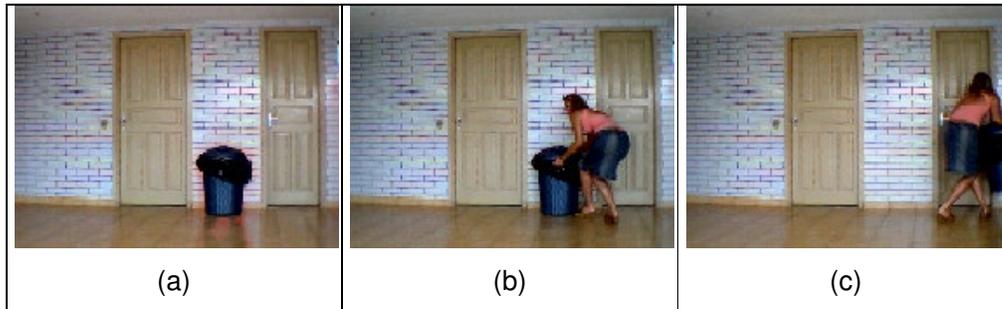


Figura 17: Seqüência para análise da atualização do fundo nos casos de alterações permanentes: (a) ambiente 4 sem ninguém; (b) pessoa entra e pega um objeto; (c) pessoa se retira.

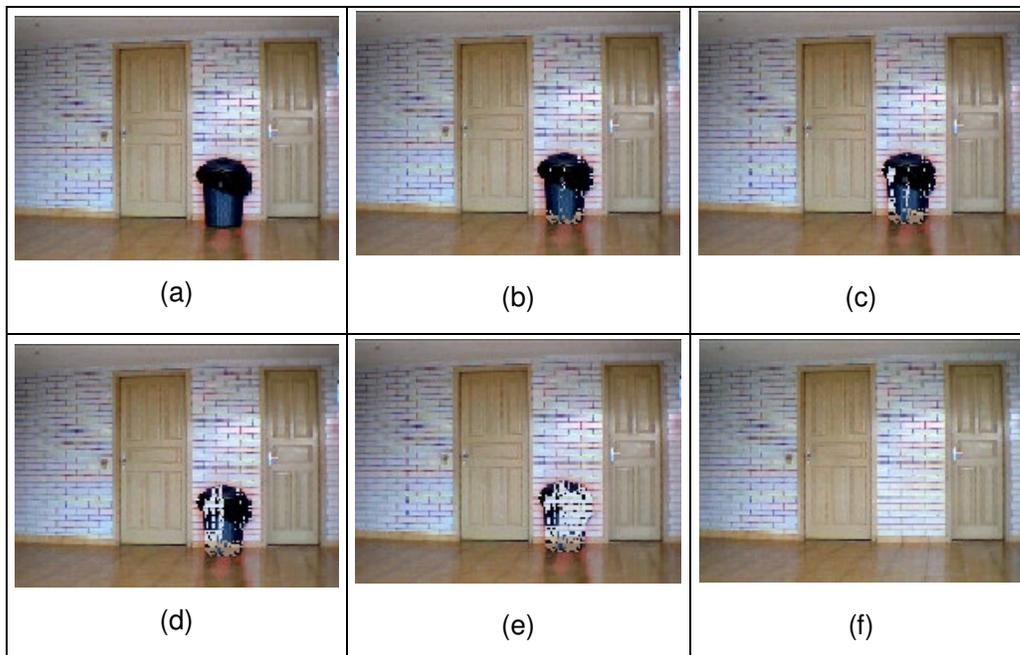


Figura 18: Atualização da estimativa de fundo relativa a alteração permanente da Figura 17: (a) lixeira pertence à estimativa; (b), (c), (d) e (e) lixeira aos poucos vai sendo eliminada da estimativa; (f) lixeira é totalmente eliminada da estimativa.

A Figura 19 ilustra mais três situações elaboradas, uma no ambiente 1, outra no ambiente 2 e outra no ambiente 3. Tais seqüências também consistiam de três momentos como no caso citado anteriormente. Ou seja, o ambiente inicia sem qualquer pessoa (Figura 19a, Figura 19b e Figura 19c), em seguida uma

pessoa entra e retira (Figura 19d) ou deixa um objeto (Figura 19e e Figura 19f) e depois deixa o local (Figura 19g, Figura 19h e Figura 19i).

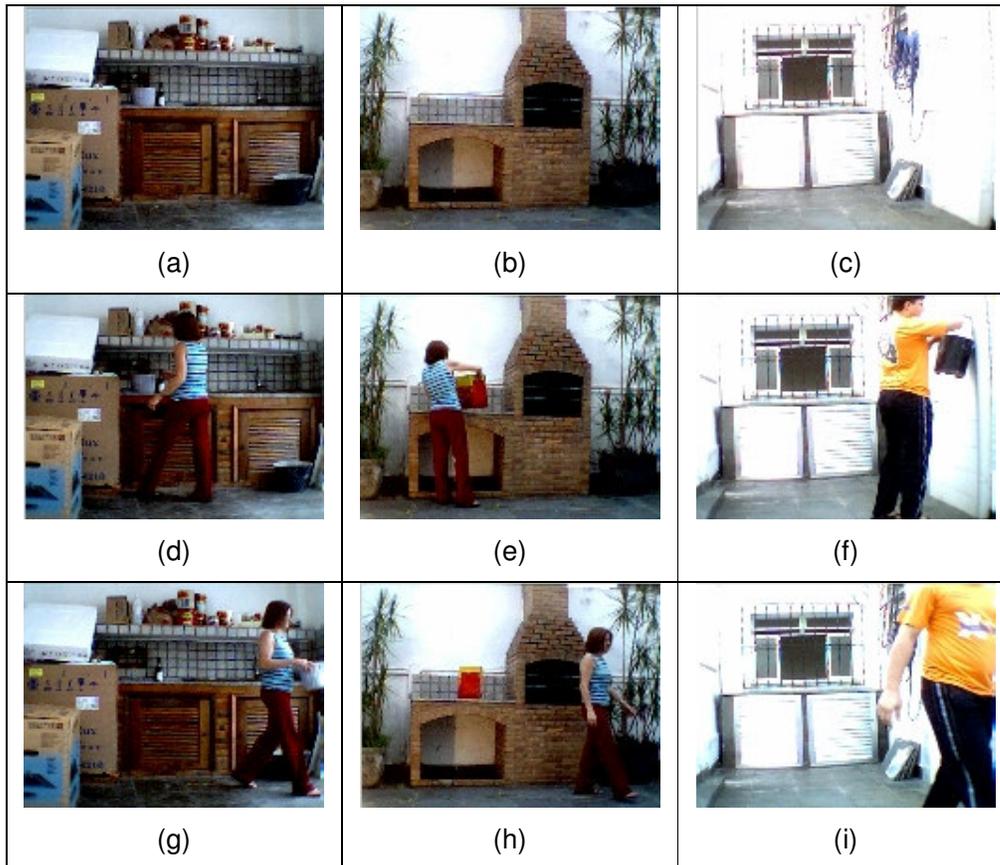


Figura 19: Seqüências para análise da atualização do fundo nos casos de alterações permanentes: ambientes 1 (a), 2 (b) e 3 (c) sem ninguém; pessoa entra e pega (d) ou deixa (e) (f) um objeto; pessoa se retira (g) (h) (i).

Assim como na primeira situação apresentada, após um determinado tempo, as alterações permanentes introduzidas nestes ambientes também foram incorporadas na estimativa de fundo através do algoritmo de atualização.

As quatro seqüências utilizadas para a avaliação deste tópico duraram 300 quadros, isto é, 1 minuto e 15 segundos. Em todas elas, as alterações simuladas foram incluídas na estimativa de fundo após 240 quadros (1 minuto), valor utilizado como  $K_{máx}$ .

#### 4.2.3. Localização / Contagem de Pessoas

Para a avaliação da localização / contagem de pessoas, além de situações simples em que pessoas circulam isoladamente no ambiente, foram

consideradas também situações envolvendo tanto oclusões parciais, quanto formação e separação de grupos. Entenda-se **grupo** como sendo um conjunto de pessoas que se tocam ou cujas imagens se sobrepõem. Desta forma, a formação de um grupo consiste em pessoas isoladas que se encontram e se tocam ou se sobrepõem, ao passo que a separação de um grupo consiste no afastamento de pessoas que estão juntas e encostadas ou sobrepostas umas às outras.

A seguir estão apresentados os procedimentos adotados para a avaliação experimental destes itens e os respectivos resultados. As métricas de desempenho nestes casos são as que se seguem:

- *Erro Médio* ( $E_{Médio}$ ): dado pela média das diferenças entre o número real de pessoas ( $p_i$ ) em cada quadro e o número encontrado pelo método ( $x_i$ ).

$$E_{Médio} = \frac{\sum_{i=1}^n (p_i - x_i)}{n}, \text{ onde } n \text{ é o número de quadros.}$$

- *Desvio Padrão* ( $D_{Padrão}$ ): dado pelo desvio da diferença acima citada ( $p_i - x_i$ ) em relação ao erro médio ( $E_{Médio}$ ) ao longo dos quadros da seqüência.

$$D_{Padrão} = \sqrt{\frac{\sum_{i=1}^n [(p_i - x_i) - E_{Médio}]^2}{n - 1}}, \text{ onde } n \text{ é o número de quadros.}$$

Cabe observar que, algumas vezes, o erro médio encontrado pode ser menor do que zero, pois a diferença entre o número real de pessoas em cada quadro e o número encontrado pelo método pode ser negativa, o que indica que o sistema errou para menos. Do contrário, quando a diferença é positiva, significa que o sistema errou para mais.

Estima-se que o número de pessoas erradas encontradas esteja dentro do intervalo cujo limite mínimo é o valor do erro médio menos o valor do desvio padrão e o limite máximo, o valor do erro médio mais o valor do desvio padrão. Como nas filmagens realizadas, o número de pessoas em cena não atinge um valor muito alto, variando de 0 a 4 pessoas apenas, é esperado que este intervalo tenha limite mínimo próximo de 0,5 e limite máximo próximo de 1,5, de forma que o sistema erre em torno de 1 pessoa.

#### 4.2.3.1. Oclusões Parciais

Para a avaliação do desempenho do sistema em situações de oclusões parciais, foi feita uma filmagem para cada um dos ambientes 1, 2 e 3, conforme ilustra a Figura 20. Primeiramente, o ambiente se encontra sem ninguém (Figura 20a, Figura 20b e Figura 20c), em seguida uma pessoa entra na cena e circula pelo ambiente (Figura 20d, Figura 20e e Figura 20f), sendo que, durante este período de movimentação pela cena, oclusões parciais são presenciadas (Figura 20g, Figura 20h e Figura 20i). Por fim, a pessoa se retira do ambiente (Figura 20j, Figura 20l e Figura 20m).

No ambiente 1, a filmagem durou 190 quadros (47,5 segundos), ao passo que nos ambientes 2 e 3, as filmagens foram de 240 quadros (1 minuto). A Tabela 4 mostra por quantos quadros houve ou não oclusões.

Tabela 4: Relação entre o número de pessoas e a quantidade de quadros das seqüências de teste de localização / contagem de pessoas nos casos de oclusões.

Número de Pessoas	Quantidade de Quadros		
	<i>Ambiente 1</i>	<i>Ambiente 2</i>	<i>Ambiente 3</i>
0	13	36	50
1 (sem oclusão)	73	101	100
1 (com oclusão)	104	103	90

O erro médio e o desvio padrão destas simulações são mostrados na Tabela 5. Pode-se observar que para o ambiente 2 os resultados não foram muito bons, pois o erro chegou a praticamente 2 pessoas a mais, o que não é adequado, pois apenas 1 pessoa circula no ambiente. Já no caso do ambiente 1, o resultado melhorou, uma vez que o sistema encontrou em torno de 1 pessoa a mais e não 2 como no ambiente 2. No ambiente 3, os resultados foram bem melhores. Pelo erro médio e desvio padrão obtidos, pode-se inferir que o sistema errou muito pouco.

Tabela 5: Erro médio e desvio padrão encontrados nas seqüências de teste de localização / contagem de pessoas nos casos de oclusões.

	<b>Ambiente 1</b>	<b>Ambiente 2</b>	<b>Ambiente 3</b>
<b><i>Erro Médio</i></b>	0,36	0,95	0,11
<b><i>Desvio Padrão</i></b>	0,67	0,90	0,31



Figura 20: Seqüências para análise da localização / contagem de pessoas nos casos de oclusões: ambiente 1 (a), 2 (b) e 3 (c) sem ninguém; pessoa entra e circula (d) (e) (f); pessoa sofre oclusão (g) (h) (i); pessoa se retira (j) (l) (m).

A discussão acerca dos motivos que levaram a resultados não tão bons em alguns casos será feita, juntamente com a discussão relativa à análise de grupos, após a apresentação dos experimentos com seqüências de vídeo de situações reais.

#### 4.2.3.2. Grupos

Para a avaliação da capacidade do sistema de contar pessoas quando elas circulam isoladamente, formam ou separam grupos, foram elaboradas três

tipos de seqüência que são descritas a seguir e ilustradas nas Figuras Figura 21, Figura 22 e Figura 23.

- i) *Pessoas isoladas* (Figura 21): o ambiente inicia sem ninguém (Figura 21a) e, em seguida, entram 3 ou 4 pessoas separadamente que circulam sem se tocar (Figura 21b). Depois as pessoas vão saindo uma a uma da cena (Figura 21c).
- ii) *Formação de grupos* (Figura 22): o ambiente inicia sem ninguém (Figura 22a, Figura 22b e Figura 22c) e, em seguida, entram 2, 3 ou 4 pessoas separadamente (Figura 22d, Figura 22e e Figura 22f). As pessoas circulam e formam grupos de 2, 3 ou 4 pessoas conforme o caso (Figura 22g, Figura 22h e Figura 22i). Depois de formados os grupos, as pessoas saem de cena, ainda em grupos, isto é, mantendo o contado estabelecido entre elas (Figura 22j, Figura 22l e Figura 22m).
- iii) *Separação de grupos* (Figura 23): o ambiente inicia sem ninguém (Figura 23a, Figura 23b e Figura 23c) e, em seguida, entram grupos de 2, 3 ou 4 pessoas (Figura 23d, Figura 23e e Figura 23f). Os grupos circulam e se separam, ou seja, deixa de existir o contato que havia entre as pessoas, quando elas entraram no ambiente (Figura 23g, Figura 23h e Figura 23i). Depois as pessoas saem de cena, ainda sem se tocar (Figura 23j, Figura 23l e Figura 23m).

Para cada um dos ambientes 1, 2 e 3, foram filmadas as seguintes seqüências:

- 1 de pessoas isoladas.
- 3 de formação de grupos (uma com 2 pessoas, uma com 3 pessoas e outra com 4 pessoas).
- 3 de separação de grupos (uma com 2 pessoas, uma com 3 pessoas e outra com 4 pessoas).

As filmagens de pessoas isoladas duraram 300 quadros (1 minuto e 15 segundos), as de formação de grupos, 160 quadros (40 segundos) e as de separação de grupos, 160 quadros (40 segundos) para os casos de 2 pessoas e 240 quadros (1 minuto) para os casos de 3 e 4 pessoas. A Tabela 6, a

Tabela 7 e a Tabela 8 apresentam a quantidade de quadros das seqüências de acordo com o número de pessoas na cena.

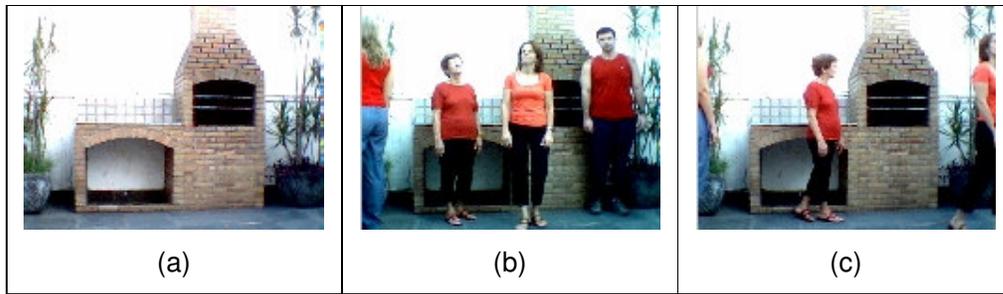


Figura 21: Seqüência para análise da localização / contagem de pessoas nos casos de pessoas isoladas: (a) ambiente 2 sem ninguém; (b) pessoas entram em cena isoladamente e circulam pelo ambiente; (c) pessoas saem sem se tocar.

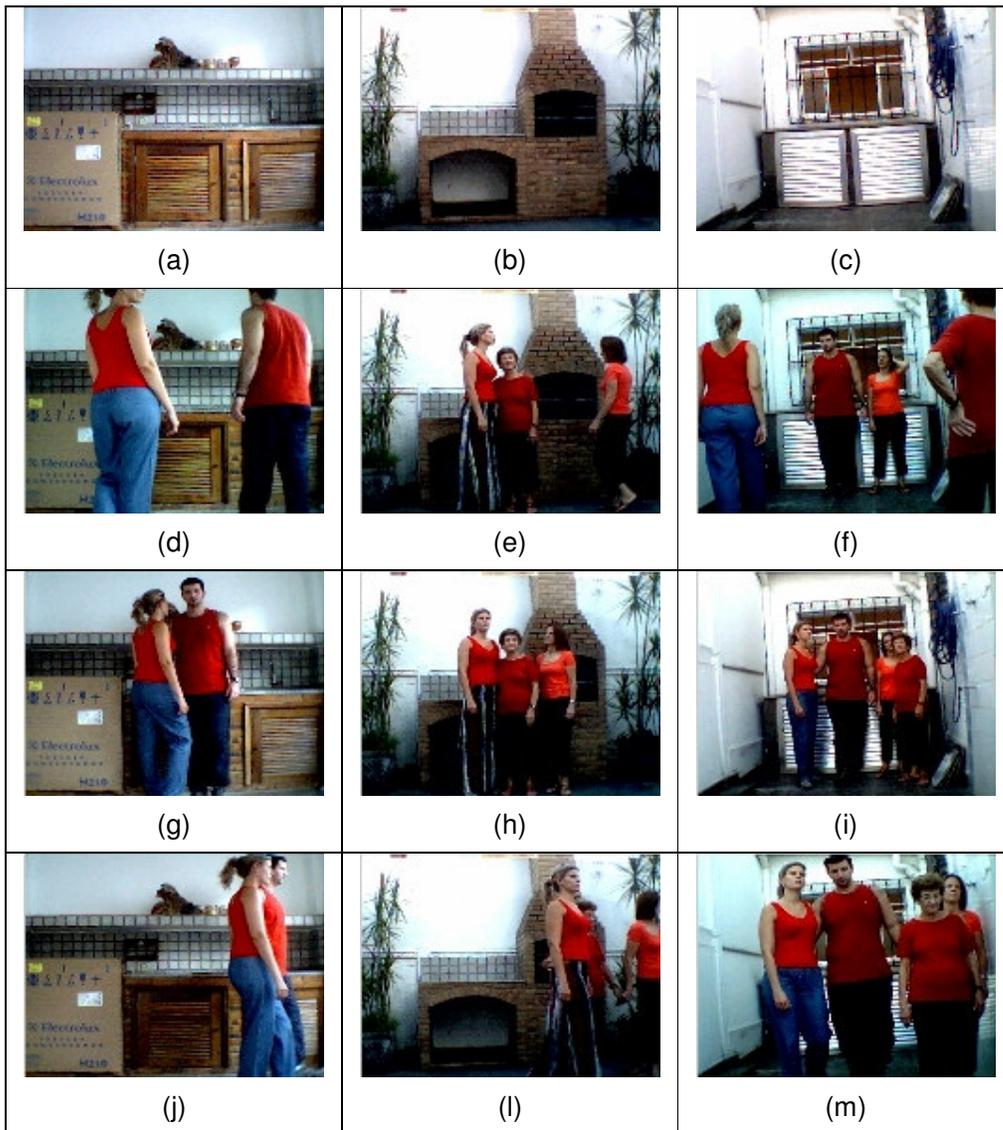


Figura 22: Seqüências para análise da localização / contagem de pessoas nos casos de formação de grupos: ambiente 1 (a), 2 (b) e 3 (c) sem ninguém; 2 (d), 3 (e) ou 4 (f) pessoas entram isoladamente na cena; grupos de 2 (g), 3 (h) ou 4 (i) pessoas se formam; grupos de 2 (j), 3 (l) ou 4 (m) pessoas se retiram.

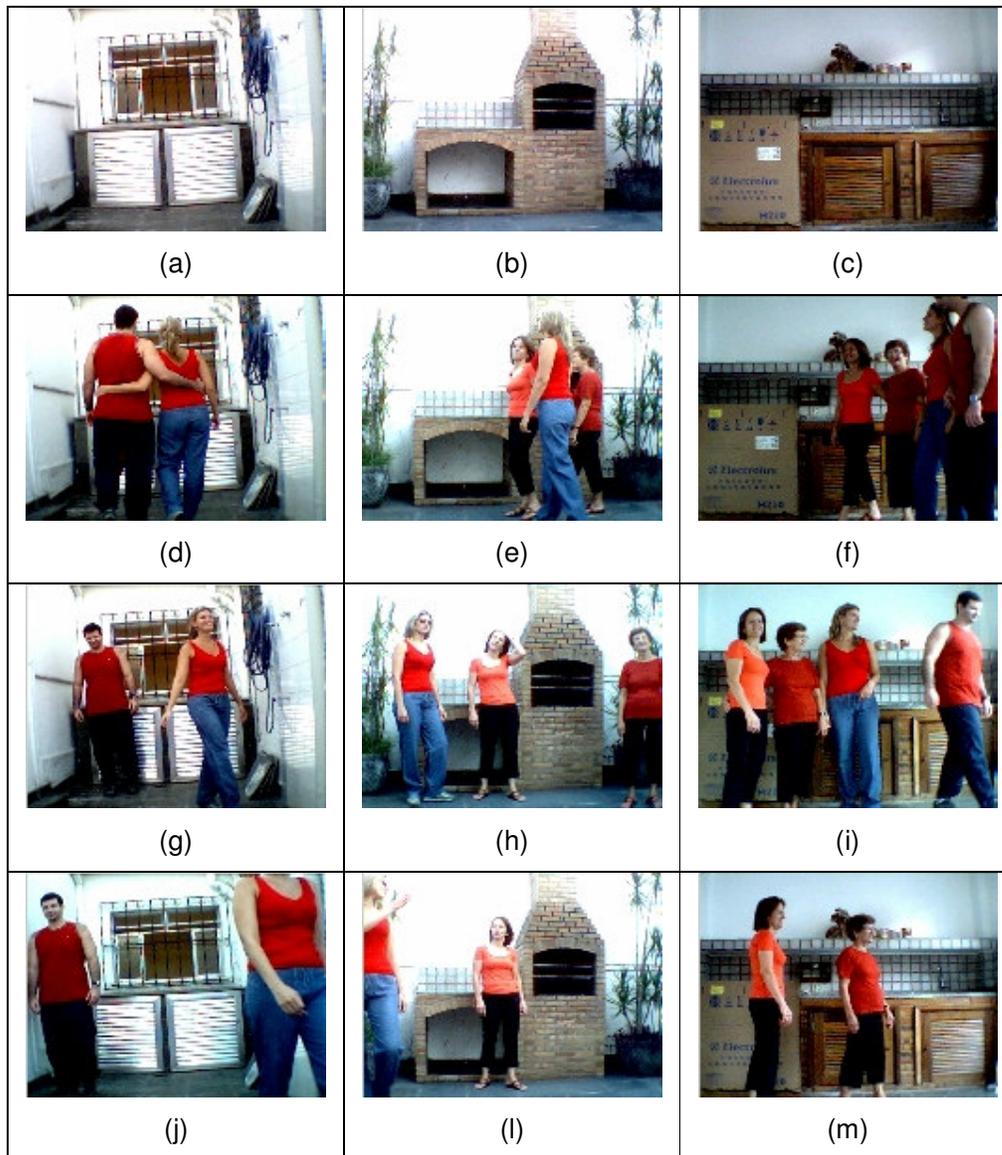


Figura 23: Seqüências para análise da localização / contagem de pessoas nos casos de separação de grupos: ambiente 1 (a), 2 (b) e 3 (c) sem ninguém; grupos de 2 (d), 3 (e) ou 4 (f) pessoas entram na cena; grupos de 2 (g), 3 (h) ou 4 (i) pessoas se separam; 2 (j), 3 (l) ou 4 (m) pessoas saem isoladamente.

A Tabela 9 mostra o erro médio e o desvio padrão encontrados em cada uma das três situações de pessoas isoladas, onde o ambiente pode não ter pessoa alguma, ou pode ter de 1 a 4 pessoas circulando. O ambiente que apresentou o melhor resultado foi o ambiente 3, onde o sistema erra em menos de 1 pessoa. No caso do ambiente 2, o resultado foi satisfatório, visto que o sistema teve erro um pouco além de 1 pessoa. Porém o ambiente 1 teve um resultado pouco satisfatório, uma vez que o sistema errou além de 2 pessoas.

Tabela 6: Relação entre o número de pessoas e a quantidade de quadros das seqüências de teste de localização / contagem de pessoas nos casos de pessoas isoladas.

Número de Pessoas	Quantidade de Quadros		
	<i>Ambiente 1</i>	<i>Ambiente 2</i>	<i>Ambiente 3</i>
0	25	17	15
1	65	84	32
2	131	69	99
3	78	76	86
4	0	54	68

Tabela 7: Relação entre o número de pessoas e a quantidade de quadros das seqüências de teste de localização / contagem de pessoas nos casos de formação de grupos.

Grupos	Número de Pessoas	Quantidade de Quadros		
		<i>Ambiente 1</i>	<i>Ambiente 2</i>	<i>Ambiente 3</i>
2 pessoas	0	61	30	31
	1	8	4	22
	2	91	125	107
3 pessoas	0	48	31	18
	1	33	37	9
	2	30	31	36
	3	48	61	96
4 pessoas	0	30	23	35
	1	33	21	6
	2	23	36	43
	3	29	21	25
	4	44	59	50

Na seqüência, a Tabela 10 apresenta o erro médio e o desvio padrão de todos os 9 casos de formação de grupos de acordo com o número de pessoas presentes nos grupos. No ambiente 1, o limite superior do erro ficou em torno de 2 pessoas para todos os três casos, o que é aceitável nos casos de formação de grupos de 3 e 4 pessoas, mas não é razoável para o caso de formação de grupos de 2 pessoas.

No ambiente 2, a formação de grupos de 3 e 4 pessoas teve um erro em torno de 1,5 pessoa, erro este satisfatório. Entretanto, na formação de grupos de 2 pessoas o resultado foi muito ruim, apresentando um erro superior a 2 pessoas. Já o ambiente 3, foi o que apresentou os melhores resultados. Para a formação de grupos de 2 pessoas o erro foi menor que 1 pessoa e para os grupos de 3 e 4 pessoas, menor que 1,5 pessoa. Considerando a quantidade de pessoas em cada caso, o desempenho foi bastante satisfatório.

Tabela 8: Relação entre o número de pessoas e a quantidade de quadros das seqüências de teste de localização / contagem de pessoas nos casos de separação de grupos.

Grupos	Número de Pessoas	Quantidade de Quadros		
		Ambiente 1	Ambiente 2	Ambiente 3
2 pessoas	0	21	36	53
	1	71	41	18
	2	68	162	89
3 pessoas	0	24	41	45
	1	44	33	57
	2	55	60	21
	3	117	105	117
4 pessoas	0	28	19	20
	1	65	32	9
	2	41	34	28
	3	37	39	95
	4	69	115	88

Tabela 9: Erro médio e desvio padrão encontrados nas seqüências de teste de localização / contagem de pessoas nos casos de pessoas isoladas circulando.

	Ambiente 1	Ambiente 2	Ambiente 3
<b>Erro Médio</b>	1,20	0,42	0,22
<b>Desvio Padrão</b>	1,14	0,80	0,59

Por fim, a Tabela 11 apresenta o erro médio e o desvio padrão de todos os 8 casos de separação de grupos também de acordo com o número de pessoas presentes nos grupos. Neste caso, o melhor resultado foi o obtido no ambiente 2, onde para os grupos de 2 e 3 pessoas o erro foi em torno de 1 pessoa e para os grupos de 4 pessoas, foi em torno de 1,5 pessoa. No ambiente 3, só foram

analisadas as situações de formação de grupos de 2 e 4 pessoas. No primeiro caso, o erro foi inferior a 1 pessoa, ou seja, dentro do desejável. No segundo caso, o erro também foi satisfatório, ficando um pouco acima de 1,5 pessoa. Porém no ambiente 1, os limites superiores dos erros ficaram bem aquém do desejado. Para grupos de 2 pessoas, ficou acima de 2 pessoas e para grupos de 3 e 4 pessoas, ficou acima de 2,5 pessoas, o que é muito ruim, tendo em vista o número de pessoas em cena.

Tabela 10: Erro médio e desvio padrão encontrados nas seqüências de teste de localização / contagem de pessoas nos casos de formação de grupos de 2, 3 e 4 pessoas.

<b>Número de Pessoas</b>	<b>Ambiente 1</b>	<b>Ambiente 2</b>	<b>Ambiente 3</b>
<b>2</b>	0,79 / 1,28	1,23 / 1,21	-0,20 / 0,58
<b>3</b>	1,06 / 0,99	0,49 / 1,05	-0,67 / 0,73
<b>4</b>	0,90 / 1,28	0,00 / 1,42	0,50 / 0,96

Tabela 11: Erro médio e desvio padrão encontrados nas seqüências de teste de localização / contagem de pessoas nos casos de separação de grupos de 2, 3 e 4 pessoas.

<b>Número de Pessoas</b>	<b>Ambiente 1</b>	<b>Ambiente 2</b>	<b>Ambiente 3</b>
<b>2</b>	1,00 / 1,31	0,40 / 0,58	-0,06 / 0,79
<b>3</b>	1,32 / 1,29	-0,11 / 0,99	-
<b>4</b>	1,26 / 1,48	0,34 / 1,34	-0,63 / 1,25

Uma outra forma de avaliação, utilizando também o erro médio e o desvio padrão, mas diferente da apresentada anteriormente, foi realizada. Todas as imagens de todas as seqüências trabalhadas foram separadas em 4 grandes grupos independentemente de sua seqüência de origem. Tais grupos foram caracterizados de acordo com a quantidade de pessoas em cada quadro. Ou seja, imagens sem ninguém formaram um grupo, imagens com 1 pessoa formaram outro grupo e assim, sucessivamente, até o grupo de imagens com 4 pessoas.

Esta forma de avaliação foi útil para analisar a relação entre o erro cometido pelo sistema e o número de pessoas presentes em cena. A Tabela 12 apresenta os resultados obtidos considerando este critério.

Para nenhuma ou 1 pessoa o erro ficou em torno de 1 pessoa, o que é aceitável para o número de pessoas em cena. Já para 2 ou 3 pessoas o limite superior do erro foi em torno de 2 pessoas, sendo ainda alto para a quantidade de pessoas. Até aqui se observa uma coerência do número de pessoas em cena com o erro encontrado, uma vez que o erro aumentou com o aumento do número de pessoas. Entretanto, nos quadros com 4 pessoas o erro caiu, ficando em torno de 1,5 pessoa, sendo inclusive tolerável. Com isto, não se tem dados suficientes para relacionar o erro encontrado com o número de pessoas.

Tabela 12: Erro médio e desvio padrão de acordo com o número de pessoas em cena.

<b>Número de Pessoas</b>	<b>Erro Médio / Desvio Padrão</b>
<b>0</b>	0,33 / 0,65
<b>1</b>	0,49 / 0,80
<b>2</b>	0,75 / 1,16
<b>3</b>	0,59 / 1,57
<b>4</b>	-0,16 / 1,45

#### **4.2.4. Detecção de Atitudes Suspeitas**

Para a avaliação de detecção de atitudes suspeitas foram simuladas 10 situações distintas nos ambientes 1, 2 e 3. Na Figura 24 imagens de 3 das simulações realizadas ilustram algumas das atitudes elaboradas. Inicialmente, o ambiente se encontra sem nenhuma pessoa (Figura 24a, Figura 24b e Figura 24c). Em seguida, uma pessoa entra e pega (Figura 24e e Figura 24f) ou deixa (Figura 24d) algum objeto no fundo e se retira do ambiente levando o objeto retirado (Figura 24h e Figura 24i) ou sem objeto algum (Figura 24g).

O procedimento para detecção de atitudes suspeitas é basicamente o utilizado para a absorção de alterações permanentes do fundo. A diferença está apenas nos tempos envolvidos.

As seqüências para avaliação de detecção de atitudes suspeitas duraram 300 quadros, isto é, 1 minuto e 15 segundos. Em todas elas, após 200 quadros (50 segundos), valor utilizado como  $K'_{máx}$ , o sistema foi capaz de detectar as atitudes suspeitas e uma indicação de alarme foi apresentada na tela do computador.

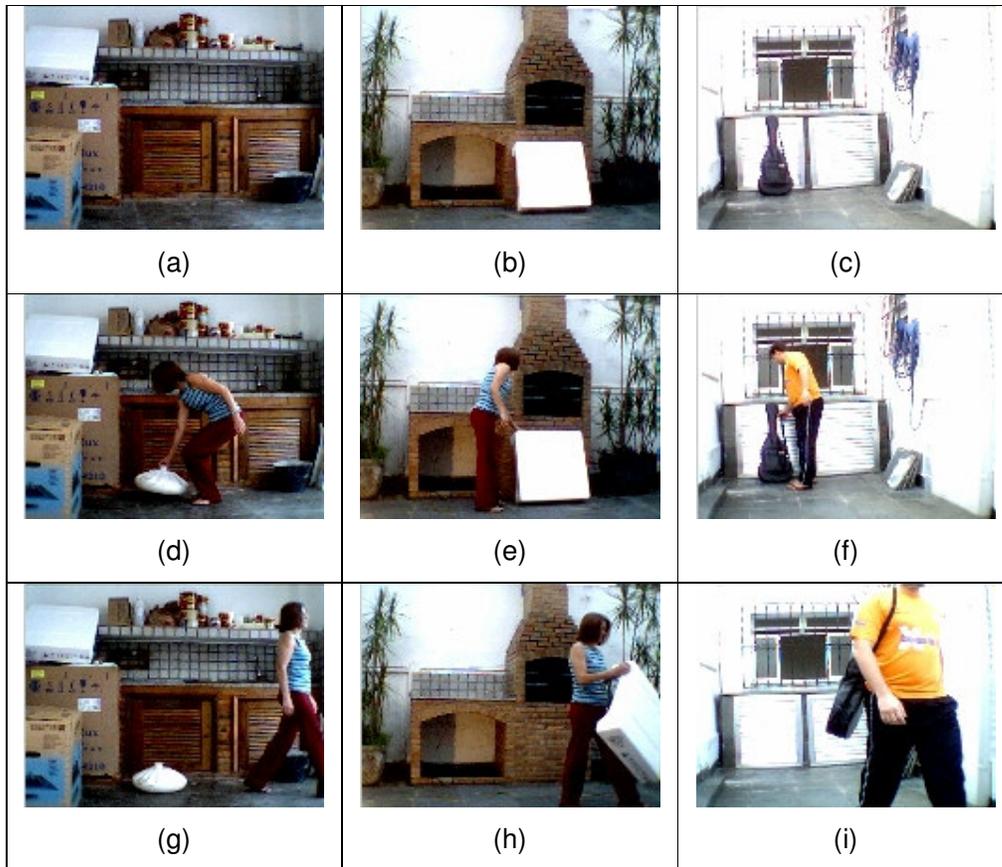


Figura 24: Seqüências para análise da detecção de atitudes suspeitas: ambiente 1 (a), 2 (b) e 3 (c) sem ninguém; pessoa entra e deixa (d) ou pega (e) (f) um objeto; pessoa se retira (g) (h) (i).

### 4.3. Experimentos com Seqüências de Vídeo de Situações Reais

Uma avaliação do sistema para situações reais foi feita com base em uma filmagem realizada em uma banca de jornal. A iluminação do ambiente e a movimentação das pessoas são espontâneas, ou seja, não há controle algum sobre elas. Deste modo o sistema pode ser testado para uma situação completamente real, mais condizente com as situações onde poderá ser implantado.

Esta filmagem teve 300 quadros de duração e foi composta por diversos momentos, conforme pode ser visto na Figura 25. Inicialmente, apenas 2 adultos isolados são identificadas em locais bem separados do ambiente (Figura 25a), uma do lado esquerdo da banca e outra no canto inferior direito. Em seguida, uma criança entra na banca (Figura 25b) e logo depois um terceiro adulto também (Figura 25c), totalizando, assim, 4 pessoas na cena, sendo que 3 destas

interagem entre si. Na seqüência, o último adulto a entrar, sai (Figura 25d) e um dos adultos que estava inicialmente na cena também (Figura 25e). A criança passa por trás do adulto que ainda permanece com ela (Figura 25f) e atravessa para o outro lado da banca (Figura 25g). Por fim, o adulto também atravessa para o outro lado da banca (Figura 25h) e se agacha, encobrendo a criança (Figura 25i).

Nesta seqüência de vídeos, além de situações de pessoas isoladas, como na Figura 25a, por exemplo, configuram-se, ainda, casos de formação e separação de grupos e oclusões, como podem ser vistos nos exemplos da Figura 25b, Figura 25d e Figura 25f, respectivamente.



Figura 25: Seqüência para análise de uma situação real: (a) 2 adultos estão presentes na cena; (b) uma criança entra; (c) um outro adulto entra; (d) um adulto sai; (e) outro adulto sai; (f) a criança passa por trás do último adulto em cena e (g) segue para o outro lado da banca; (h) o adulto também segue para o outro lado da banca; (i) o adulto se agacha e encobre a criança.

A Tabela 13 mostra a quantidade de quadros de acordo com o número de pessoas em cada cena. Em seguida, a Tabela 14 apresenta o erro médio e o desvio padrão encontrados neste caso.

Tabela 13: Relação entre o número de pessoas e a quantidade de quadros da seqüência de teste de uma situação real.

<b>Número de Pessoas</b>	<b>Quantidade de Quadros</b>
1	117
2	69
3	38
4	76

Tabela 14: Erro médio e desvio padrão encontrados na seqüência de teste de uma situação real.

	<b>Situação Real</b>
<b><i>Erro Médio</i></b>	0,88
<b><i>Desvio Padrão</i></b>	1,02

Para esta seqüência os resultados podiam se melhores, pois o limite superior do erro foi de praticamente 2 pessoas, o que não é adequado para uma seqüência de imagens, onde se têm no máximo 4 pessoas em cena.

Pode-se observar que tanto os resultados obtidos nesta situação real quanto os obtidos na avaliação da localização / contagem de pessoas com vídeos simulados não apresentaram resultados satisfatórios o tempo todo. Em alguns casos os resultados atenderam bem ao esperado, em outros nem tanto. Por outro lado, os resultados obtidos na avaliação da separação de fundo / primeiro plano, da atualização da estimativa de fundo e da detecção de atitudes suspeitas apresentaram resultados muitos bons, atendendo ao esperado em todos os casos.

Na verdade, 3 problemas, que estão levando a estes erros maiores que o desejável, podem ser identificados na fase de localização / contagem de pessoas. O primeiro deles diz respeito à tolerância na coerência de movimento que não pode ser muito grande tendo em vista que os movimentos das pessoas são completamente aleatórios. Desta forma, uma tolerância baixa pode levar ao agrupamento de segmentos pertencentes a duas pessoas distintas, mas que são adjacentes e dados como coerentes.

Outro problema se refere a segmentos que foram rotulados de forma diferente indevidamente e não são agrupados novamente, ainda que pertençam a uma só pessoa, pois o método entende que rótulos diferentes representam pessoas diferentes. O terceiro e último problema nada mais é do que uma razão para a rotulação indevida citada no problema anterior. Algumas vezes o procedimento utilizado para a remoção de sombras elimina partes do corpo de uma pessoa que têm pouco contraste em relação ao fundo. Esta eliminação pode causar a separação de uma única pessoa em duas partes distintas que são rotuladas de maneira distinta. Com isto recai-se no segundo problema citado. As sugestões para o tratamento destes itens serão discutidas nos comentários finais.

#### 4.4. Sensibilidade aos Valores dos Parâmetros

Nesta seção discute-se a sensibilidade do sistema aos valores dos parâmetros do método proposto. O significado de cada um destes parâmetros, a forma como eles foram encontrados e os valores assumidos para este trabalho também são apresentados.

A Tabela 15 apresenta a descrição e os valores dos parâmetros fixos do sistema, enquanto que a Tabela 16 apresenta as mesmas informações só que relativas aos parâmetros ajustáveis do sistema.

Tabela 15: Descrição e valores dos parâmetros fixos do sistema.

$A_{máx}$	Define o tamanho máximo para cada segmento originado na segmentação watershed.	100
$A_{mín}$	Define o tamanho mínimo para cada segmento originado na segmentação watershed.	30
$H$	Define os marcadores utilizados na segmentação watershed.	0.1

Os valores dos parâmetros  $T_L$ ,  $T_H$  e  $T_C$  apresentados para cada ambiente são adequados para o correto funcionamento do sistema, incluindo o tratamento de sombras. Entretanto, na análise exclusiva do tratamento de sombras apresentado na seção 4.2.1.1, os valores utilizados para estes parâmetros foram um pouco mais tolerantes, uma vez que simulou-se sombras mais fortes do que

as usualmente encontradas em ambientes razoavelmente bem iluminados. Neste caso, os valores utilizados foram  $T_L = 0.1$ ,  $T_H = 0.5$  e  $T_C = 0.4$  para os 3 ambientes analisados, 1, 2 e 3.

Tabela 16: Descrição e valores dos parâmetros ajustáveis do sistema.

$T_L$	Define um primeiro patamar de similaridade entre a imagem e o fundo.	<i>Ambiente 1</i>	0.1
		<i>Ambientes 2 e 3</i>	0.1
		<i>Ambientes 4 e 5</i>	0.08
$T_H$	Define um segundo patamar de similaridade entre a imagem e o fundo.	<i>Ambiente 1</i>	0.3
		<i>Ambientes 2 e 3</i>	0.4
		<i>Ambientes 4 e 5</i>	0.2
$T_C$	Caracteriza os pixels relativos a sombras.	<i>Ambiente 1</i>	0.15
		<i>Ambientes 2 e 3</i>	0.3
		<i>Ambientes 4 e 5</i>	0.09
$K_{m\acute{a}x}$	Define o número de quadros consecutivos máximo permitido para que mudanças no fundo não sejam incluídas na sua atualização.	240	
$p$	Define o grau de influência de um pixel dado como fundo na sua estimativa.	0.01	
$e_b$	Representa o quão similar é um pixel do primeiro plano do quadro atual em relação ao quadro anterior.	0.35	
$R_{m\acute{a}x}$	Define o deslocamento máximo que uma região de interesse pode sofrer entre dois quadros consecutivos.	4	
$\pi_{max}$	Define o percentil máximo admitido para considerar que uma região correspondente foi encontrada.	0.45	
$D_{m\acute{a}x}$	Define a distância máxima tolerada para considerar dois vetores como coerentes.	12	
$S_{m\acute{i}n}$	Define a área mínima estabelecida para manter na análise um segmento isolado.	130	
$K'_{m\acute{a}x}$	Define o número de quadros consecutivos máximo para que uma alteração no fundo não constate uma atitude suspeita.	200	

Todos estes parâmetros são definidos experimentalmente, entretanto existem diferenças na definição deles.  $T_L$ ,  $T_H$ ,  $T_C$ ,  $e_b$  e  $H$  são encontrados de forma mais subjetiva, pois expressam valores com significados implícitos. Por outro lado,  $\rho$ ,  $K_{m\acute{a}x}$ ,  $R_{m\acute{a}x}$ ,  $\pi_{m\acute{a}x}$ ,  $D_{m\acute{a}x}$ ,  $S_{m\acute{a}x}$ ,  $K'_{m\acute{a}x}$ ,  $A_{m\acute{a}x}$  e  $A_{m\acute{i}n}$  são mais objetivos, pois representam valores com significados explícitos, sendo, desta maneira, mais simples de serem obtidos. Para os parâmetros onde cabe a análise de sensibilidade ( $T_L$ ,  $T_H$ ,  $T_C$ ,  $e_b$ ,  $V_{m\acute{a}x}$  e  $D_{m\acute{a}x}$ ), foram variados seus respectivos valores em  $\pm 10\%$ , sendo que o desempenho do sistema não foi afetado significativamente.

Apesar dos parâmetros  $T_L$ ,  $T_H$  e  $T_C$  não se apresentarem sensíveis em relação a uma variação de  $\pm 10\%$ , eles tiveram de ser ajustados para ambientes distintos. Isto ocorre, pois os ambientes apresentam características e condições diferentes, de forma que os parâmetros precisam de adequação. Esta adequação não altera muito os valores de forma absoluta, entretanto, percentualmente, a variação é superior a 10% (variação tolerável para a análise de sensibilidade).