

## 2 Abordagens Anteriores

Com base na estrutura planejada para o desenvolvimento do sistema em questão, diversos trabalhos anteriores, envolvendo esta área, foram analisados. Alguns são mais abrangentes e incorporam a maior parte das etapas indicadas no capítulo anterior, outros tratam de forma mais específica apenas uma ou duas etapas. Além disto, em alguns casos as etapas se confundem, isto é, num único algoritmo se obtém o resultado pretendido em duas ou três etapas conjuntamente.

Os sistemas analisados são bastante diversos, abordando casos de câmeras múltiplas ou de uma única câmera, imagens a cores ou em preto e branco e estáticas ou dinâmicas. No que se refere à funcionalidade do sistema, pode-se citar casos de tratamento de pessoas isoladas ou múltiplas, de oclusões, de localização ou não de partes do corpo e de definição da trajetória das pessoas.

A Figura 1 apresenta a estrutura do sistema proposto neste trabalho. Em seguida, cada uma das tarefas que ele deve realizar é comentada.

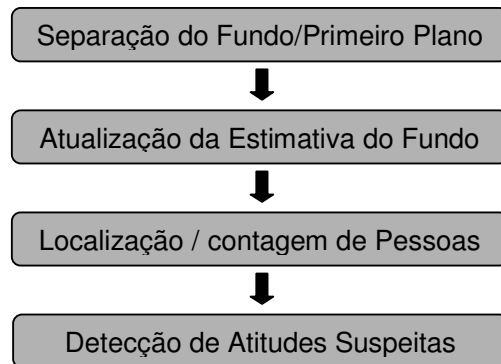


Figura 1: Estrutura básica do sistema proposto.

O primeiro passo realizado pelo sistema proposto é a separação do fundo / primeiro plano. Na prática este passo deve produzir uma matriz lógica com o mesmo número de linhas e colunas da imagem de entrada, que indica quais pixels da imagem de entrada contém informação relativa ao fundo (matriz = "verdadeiro") ou ao primeiro plano (matriz = "falso"). Desta forma, tudo que

estiver no primeiro plano, ou seja, tudo que for passível de representar uma pessoa, será representado pelo bit “0”, ao passo que o restante, relativo ao fundo da imagem, será representado pelo bit “1”.

Um método simplificado para produzir esta matriz consiste em fazer a diferença entre a imagem de entrada e uma estimativa do fundo, obtida de algum modo. Pixels em que a disparidade entre imagem e estimativa do fundo for elevada serão “falsos” (pertencem ao primeiro plano), caso contrário serão “verdadeiros” (pertencem ao fundo). Nesta etapa, admite-se, portanto, que uma estimativa de fundo está disponível. A etapa seguinte trata exatamente da atualização desta estimativa a cada quadro, de maneira que variações de iluminação ao longo do tempo ou mudanças permanentes no fundo sejam absorvidas pela estimativa.

O próximo passo é a localização / contagem de pessoas dentro do que se identificou como primeiro plano. Nesta etapa é feita a análise de grupos e de oclusões parciais. Por fim, uma análise das atualizações do fundo permite que sejam detectadas mudanças permanentes no fundo que podem indicar o roubo de um objeto ou um objeto que foi deixado no ambiente.

As seções a seguir se referem a cada uma das etapas mencionadas, com exceção da última que não é abordada nestes artigos, e apresenta as descrições resumidas dos métodos pesquisados para a implementação de cada uma delas.

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

O procedimento de separação de fundo / primeiro plano deve atender os seguintes requisitos.

- i) Eliminação de ruídos.
- ii) Tratamento de sombras.

Uma das maneiras mais simples de segmentar uma imagem, isto é, de destacar as regiões relativas ao primeiro plano, é através da subtração entre a imagem de entrada e um modelo existente de fundo, tal como em [CAI 95], [KETTNAK 99] e [ATSUSHI 02]. A obtenção e / ou atualização deste modelo de fundo pode ser feita de diversas maneiras conforme será explicitado na seção seguinte, porém, independente de como o modelo for obtido, o processo de subtração é basicamente o mesmo. A separação entre fundo e primeiro plano é feita neste caso, através de uma limiarização aplicada ao módulo da diferença.

Nem todos os autores pesquisados, que possuem sistemas com uma etapa relativa à separação do fundo / primeiro plano, abordam os itens i) e ii) acima e, com exceção de Haritaoglu [HARITA 00], todos utilizam, de alguma forma, a **subtração de fundo**. Os métodos utilizados por estes autores estão apresentados a seguir com destaque para suas especificidades.

Os trabalhos descritos em [PIAU 02], [LU 01], [WOJTASZ 02], [SEGEN 96], [KHAN 01], [TAI 04] e [KUMAR 03] têm em comum não só a subtração de fundo, mas também a **limiarização** da matriz obtida após tal subtração. Detalhes presentes em cada um destes trabalhos são importantes em termos da eficiência de cada sistema, uma vez que ruídos (i) e sombras (ii) podem ou não ser tratados.

Piau [PIAU 02] faz a subtração do fundo e, em seguida, limiariza apenas as componentes R e G da diferença em função de suas respectivas variâncias, pois desta forma constatou evitar a inclusão de sombras (ii) e de ruídos (i) na análise. Já Wojtaszek [WOJTASZ 02] aplica um filtro mediana à matriz obtida na subtração de fundo, antes de efetuar a limiarização, de forma a eliminar possíveis ruídos (i). Além disto, uma operação morfológica de fechamento é implementada para remover pequenos fragmentos.

Lu [LU 01], após a subtração do fundo, utiliza a média e o desvio padrão da imagem de fundo para limiarizar a matriz diferença obtida e, assim como Khan [KHAN 01], faz uma filtragem morfológica da matriz lógica através de dilatações e erosões para eliminar ruídos (i) e suavizar o contorno das pessoas. Tai [TAI 04] utiliza um limiar variável para a obtenção da máscara binária. Tal limiar é obtido de acordo com a distribuição de intensidade em uma determinada região da imagem, encontrada após a subtração de fundo, onde não existem objetos relativos ao primeiro plano. Este procedimento confere ao sistema robustez contra variações de iluminação. Além disto, para que sejam eliminados ruídos do tipo “sal e pimenta” (i), é utilizado após este procedimento um filtro espacial horizontal. Um retângulo de tamanho 5x1 passa sobre cada pixel da imagem e verifica-se se na região sobre a qual está passando, mais de 2 pixels estão acesos, neste caso, o pixel em questão deve ser igualado a “1”, caso contrário, a “0”.

Kumar [KUMAR 03], baseando-se no trabalho de Prati [PRATI 01], procura fazer uma limiarização que detecta sombras (ii), através do sistema de coordenadas HSV (hue - matiz, saturation - saturação, value - brilho) [GONZALE 02]. Este trabalho mostra que os valores de brilho dos pixels relativos a sombras são sempre menores que a média dos seus respectivos valores do modelo de

fundo e se encontram em uma faixa entre dois limiares. Por outro lado, nas componentes de crominância a diferença devido a sombras é mínima.

Huang [HUANG 03], após a subtração de fundo, não limiariza a matriz diferença, mas separa o fundo de uma forma simplificada baseada em **janelas**, que auxilia na eliminação de ruídos (i). Esta separação faz uso de uma região de 5x5 pixels que percorre toda a imagem. Se o número de pixels do primeiro plano dentro desta janela for inferior a um determinado valor, tais pixels retornam “0”, caso contrário, retornam “1”. Desta maneira, obtém-se uma imagem binária, na qual o algoritmo k-means de agrupamento é aplicado para agrupar pixels em grupos. Em seguida, os grupos encontrados são unidos em grupos maiores que representam pessoas. Para evitar dificuldades na distinção dos casos onde dois grupos representam uma só pessoa e onde um grupo representa duas pessoas, as informações contidas nos últimos 10 quadros são mantidas para maiores esclarecimentos. Butler [BUTLER 03] também faz uso de um algoritmo de agrupamento.

Shio [SHIO 91] se baseia no **movimento** para realizar a separação do fundo / primeiro plano. A partir da matriz gerada pela subtração de fundo, obtém os contornos das regiões do primeiro plano e, em seguida, uma imagem com apenas 3 valores. 1 e -1 representam regiões de movimento em duas direções distintas e 0 representa regiões relativas ao fundo ou ao contorno dos objetos em movimento. Feito isto, as regiões de movimento são divididas em sub-regiões de acordo com a direção de seus respectivos movimentos. Por fim, tais sub-regiões são unidas em conformidade com um modelo probabilístico de uma pessoa em movimento que possui informações de tamanho e movimento. As hipóteses mais prováveis são aceitas no momento do agrupamento. Assim como Hayfron [HAYFRON 01], que aplica um **operador de gradiente** para obtenção da máscara binária, Shio não trata nem ruídos (i), nem sombras (ii) especificamente. Entretanto, através da análise de movimento por ele proposta, consegue tratar casos de oclusão já nesta etapa de separação do fundo / primeiro plano.

Em [HARITA 00], Haritaoglu implementa a separação do fundo / primeiro plano da imagem sem envolver subtração de fundo. Esta separação também é feita por **limiarização** e se baseia em valores associados a cada pixel do modelo de fundo, são eles: intensidades mínima e máxima e média das maiores diferenças absolutas entre quadros. Feito isto tem-se uma máscara binária que, de maneira similar a Huang [HUANG 03], irá passar por uma eliminação de ruídos (i) baseada em regiões, seguida de uma análise de componentes

conectados para fornecer uma denominação única para cada objeto encontrado na imagem.

## 2.2. Atualização da Estimativa do Fundo

Como foi dito anteriormente, para que a separação do fundo / primeiro plano seja realizada é necessária uma estimativa de fundo. Nesta seção serão apresentados os métodos pesquisados para a obtenção deste modelo de fundo. Idealmente, tal procedimento deve:

- i) Permitir que o fundo seja modelado mesmo com a presença de pessoas.
- ii) Admitir alterações permanentes na disposição dos objetos do fundo.
- iii) Tratar variações de iluminação ambiente.

Shio [SHIO 91] trata os aspectos i) e iii) acima citados, e tem a opção de tratar também o aspecto ii). Utiliza **filtragem modal** no domínio do tempo, isto é, toma para a imagem de fundo o valor mais freqüente (moda) de cada pixel de uma seqüência de imagens, podendo esta conter pessoas. Além disto, tal seqüência é revisada periodicamente para tratar mudanças nas condições de iluminação.

Piau [PIAU 02] e Huang [HUANG 03] fazem uso de **distribuições gaussianas** para modelar o fundo, sendo que cada um desenvolve um método particular. Piau [PIAU 02] trata os itens i) e iii), porém possui características indesejadas. A modelagem do fundo é feita com base em uma função densidade de probabilidade gaussiana, a partir de uma seqüência de quadros que pode conter algumas pessoas (i). Entretanto, estas pessoas precisam ser localizadas através do movimento e retiradas da estimativa de fundo, isto é, as pessoas presentes na seqüência de imagens não são tratadas diretamente pelo método proposto, e sim por um artifício auxiliar. Logo, numa primeira análise, esta não seria a forma mais prática de se modelar fundos que contenham algumas pessoas. Para o tratamento de variações de iluminação (iii), [PIAU 02] utiliza equações que atualizam a média e a variância quadro a quadro. Além disto, o fundo é remodelado, em caso de movimentações significativas da câmera, através de um algoritmo específico que as detecta.

Huang [HUANG 03] mescla o uso de distribuições Gaussianas com um **modelo estatístico**. Pré-computa um modelo estatístico de fundo com base em 25 quadros onde variações de iluminação (iii) são simuladas para que problemas

relativos à iluminação sejam eliminados. Desta maneira, tais problemas são resolvidos de forma trabalhosa e talvez até precária, uma vez que é difícil simular todas as condições de iluminação possíveis. Com o modelo estatístico em mãos, Huang [HUANG 03] formula para cada pixel do fundo uma função de distribuição Gaussiana tridimensional relativa ao canal de cores RGB. Os pixels com maior probabilidade de pertencer ao fundo em cada uma das componentes de cor serão subtraídos da imagem.

Rossi [ROSSI 94] e Haritaoglu [HARITA 00] utilizam um procedimento que envolve a análise do **movimento dos pixels** para a modelagem do fundo. Em [ROSSI 94] o primeiro plano também é retirado de acordo com os movimentos detectados na imagem. Porém, neste caso, um histograma é elaborado através da contagem do número de pixels em movimento em cada coluna de uma determinada região de forma que picos representem pessoas em movimento e vales representem regiões referentes ao fundo. Apesar deste método permitir a presença de pessoas durante a modelagem (i) e a variação na disposição dos objetos do fundo (ii), o movimento das pessoas é restrito a uma direção, o que não atende um dos requisitos principais estabelecido para o sistema a ser construído.

Haritaoglu [HARITA 00] trata todos os itens anteriormente citados, i), ii) e iii), e busca, na atualização da estimativa de fundo, o tratamento de ambientes abertos em que ocorram pequenas movimentações. Para isto, primeiramente, um filtro mediana é aplicado aos valores dos pixels durante alguns segundos no vídeo, de forma a distinguir os pixels em movimento dos pixels estacionários. Em seguida, os pixels estacionários são processados para construir o modelo inicial de fundo onde três valores são associados a cada pixel: o valor mínimo e o valor máximo de sua intensidade e a diferença máxima entre as intensidades de quadros consecutivos observados durante o período de modelagem. Tais valores são utilizados na separação do fundo / primeiro plano, como citado na seção anterior.

Além disto, Haritaoglu [HARITA 00] atualiza o modelo de fundo para que seja dado o correto tratamento aos movimentos na imagem. Para tanto, faz uso de um algoritmo que analisa o número de vezes que um determinado pixel foi classificado como pertencente ao fundo e ao primeiro plano e o número de quadros desde a última vez que um pixel recebeu a classificação de primeiro plano. Na medida em que um pixel for classificado muitas vezes como fundo ou primeiro plano, diferentemente de sua última classificação, esta será atualizada.

Wojtaszek [WOJTASZ 02] atualiza seu modelo quadro a quadro, acompanhando variações de iluminação (iii) e alterações na disposição de objetos (ii), permitindo ainda a presença de pessoas durante a modelagem (i). Esta, por sua vez, é feita através da **soma ponderada** entre a imagem atual e a última estimativa existente do fundo. Os pesos de cada pixel dependem dos movimentos dos pixels que ocorreram da imagem anterior para a imagem atual.

Cai [CAI 95] utiliza um processo de **reconstrução dinâmica** para a obtenção do modelo do fundo. Tal processo se dá através da interseção entre as matrizes lógicas resultantes das limiarizações da diferença entre o quadro atual e o quadro anterior e da diferença entre o quadro atual e o quadro posterior. As regiões cobertas existentes nesta interseção são preenchidas, o procedimento se repete até que todo o fundo tenha sido preenchido, ou seja, reconstruído. Neste caso, apenas o item i) é abordado, isto é, a presença de pessoas.

Ramanan [RAMANAN 03] não estima o fundo, visto que não utiliza nenhum processo para correção da estimativa. Seu método busca encontrar pessoas e definir sua trajetória de forma independente do fundo, o que significa que o fundo pode inclusive conter objetos que se movimentem, como, por exemplo, carros trafegando. Os procedimentos desenvolvidos por Ramanan garantem que, mesmo em movimento, os objetos do primeiro plano não sejam confundidos com pessoas. Desta forma, elimina-se a etapa de modelagem de fundo.

Outros trabalhos que não atendem nenhum dos pré-requisitos apresentados podem ser encontrados, como Lu [LU 01] e Wren [WREN 97] que também obtêm o fundo através de **modelos estatísticos** baseados nas estimativas de covariância, média e desvio padrão de uma seqüência de imagens sem nenhuma pessoa presente. De forma mais trivial, Atsushi [ATSUSHI 02] utiliza simplesmente uma **imagem prévia** do fundo como estimativa.

### 2.3. Localização / Contagem de Pessoas

Para a localização / contagem de pessoas diversas técnicas foram propostas. Em alguns casos, esta etapa não ocorre explicitamente, uma vez que se considera que tudo aquilo que foi separado como parte do primeiro plano é relativo a pessoas. Isto ocorre em [SHIO 91], [PIAU 02], [CAI 95], [SEGEN 96], [KHAN 01] e [HUANG 03].

Uma boa localização / contagem de pessoas deve abordar os seguintes aspectos:

- i) Análise de grupos de pessoas.
- ii) Oclusões parciais.

Uma técnica muito utilizada nos métodos de localização / contagem de pessoas é a **análise de silhueta** que tem como vantagem o atendimento do item i) acima destacado. Wojtaszek [WOJTASZ 02] analisa a silhueta dos objetos procurando encontrar formas referentes a cabeças com base nos contornos existentes. Em caso positivo, observa se abaixo das cabeças encontradas existem pixels correspondentes ao corpo da pessoa. Isto se dá através da definição de uma região abaixo da suposta cabeça na qual deve haver um percentual suficiente dos pixels dados como primeiro plano. Caso este percentual seja atendido, considera-se que uma pessoa foi localizada. Desta forma eliminam-se objetos que aparentemente poderiam ser consideradas como pessoas por parecerem cabeças, mas que na verdade não são. O método tem dificuldades de lidar com situações em que as pessoas não estão em pé ou sentadas, como por exemplo, encurvadas.

Em W<sup>4</sup> [HARITA 00], Haritaoglu gera um conjunto de características relacionadas à silhueta e à aparência dos objetos encontrados no primeiro plano para poder diferenciar pessoas de objetos, localizar pessoas se movimentando em grupos (i) e, ainda, detectar pessoas carregando objetos. Para isto, utiliza características globais que envolvem a coordenada média, o eixo principal e os histogramas de projeção vertical e horizontal de cada objeto. Analisa também características locais, tais como vértices convexos e vales no contorno da silhueta, além da análise de periodicidade.

Na medida em que [HARITA 00] localiza apenas uma pessoa, o sistema Ghost [HARITA 98], também de Haritaoglu, implementa uma análise da postura do corpo desta pessoa para computar as similaridades entre os histogramas de projeção horizontal e vertical da silhueta e posturas principais pré-determinadas. A postura que tiver maior similaridade com a silhueta detectada na imagem será tomada como a postura estimada da pessoa. Em seguida, um algoritmo é aplicado para encontrar possíveis partes do corpo no contorno da silhueta. A localização da cabeça é estimada através do eixo principal, dos vértices da silhueta e também através da postura estimada. De posse desta localização, uma análise é realizada para eliminar vértices do contorno que não representam partes do corpo e para mapear aqueles que representam partes do corpo de acordo com uma ordem topológica específica.



Caso seja detectado que a pessoa está de pé, através do sistema Backpack [HARITA 99a], Haritaoglu analisa ainda se esta pessoa está carregando algum objeto ou não. Como esta diferenciação entre pessoa e objeto não é relevante para o sistema proposto nesta dissertação, maiores detalhes sobre este procedimento não serão aqui tratados.

O sistema W<sup>4</sup> [HARITA 00] de Haritaoglu é um pouco mais sofisticado que os outros aqui apresentados, uma vez que envolve a relação entre diferentes sistemas desenvolvidos separadamente por ele. Vale a pena deter-se um pouco mais em descrevê-lo. O sistema está ilustrado na Figura 2 de forma esquemática.

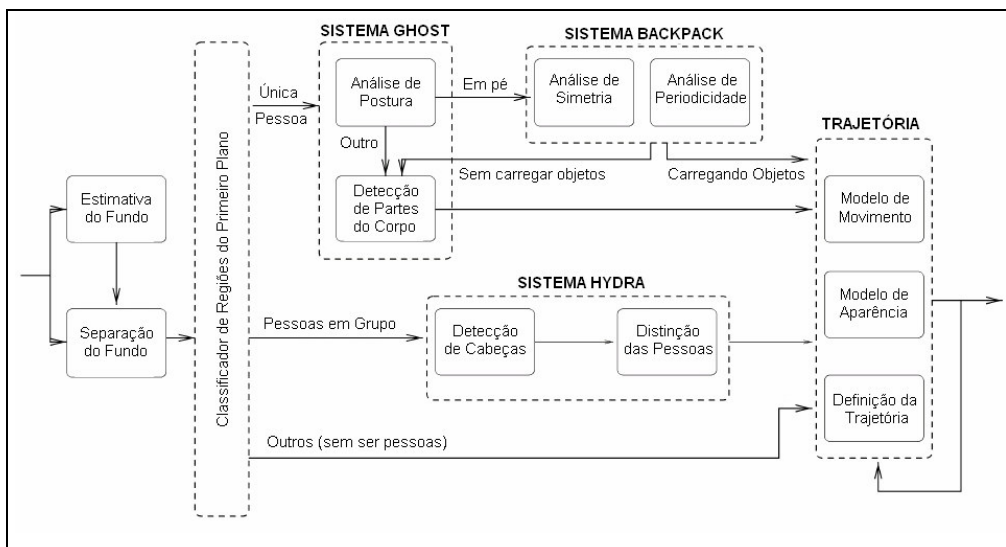


Figura 2: Sistema W<sup>4</sup> de Haritaoglu.

Como se pode notar, a localização das partes do corpo é realizada em duas situações, ou quando se observa pela análise de postura que a pessoa não está em pé, ou quando se observa pelo sistema Backpack que a pessoa está em pé e não carrega objeto algum. Em qualquer um destes casos, após a localização das partes do corpo o sistema segue para o módulo de definição da trajetória da pessoa, assim como quando a pessoa está em pé e carrega algum objeto.

Ainda restam dois outros casos a serem analisados, o primeiro quando o objeto detectado não é relativo a uma pessoa. Mesmo assim o objeto é encaminhado para o módulo de definição de trajetória. Já o segundo caso ocorre quando o objeto detectado é referente a um grupo de pessoas. Nesta situação, o sistema Hydra [HARITA 99b] se encarrega de identificar os pixels presentes

neste objeto que pertencem a cada uma das pessoas existentes no grupo. Para isto, de forma similar ao sistema Ghost [HARITA 98], Haritaoglu [HARITA 99b] busca encontrar cabeças através da análise de características globais e locais. O sistema analisa o histograma de projeção vertical da imagem e une as informações obtidas através dele com as provenientes da análise dos vértices e dos vales presentes no contorno da silhueta, eliminando, assim, detecções erradas.

Uma vez que as cabeças presentes no grupo foram identificadas, resta saber quais pixels pertencem a quais corpos. Esta última etapa do processo se dá de acordo com a distância entre tais pixels e os eixos dos troncos definidos de acordo com as cabeças encontradas. Desta forma, tem-se um objeto único separado de acordo com o número de pessoas representadas em tal objeto. Feito isto, conforme a Figura 2, a próxima etapa é a definição da trajetória que também não será abordada neste trabalho, por não fazer parte dos objetivos propostos.

Em outro sistema desenvolvido por Haritaoglu [HARITA 01] a localização / contagem de pessoas também é feita através da identificação das cabeças na silhueta. A princípio, este sistema trata de pessoas paradas, já que se refere à contagem de indivíduos na frente de quadros de aviso. Todavia, antes e depois de parar na posição onde observará um determinado aviso, a pessoa entra e sai do campo de visão do sistema. Com isto, complementa-se o sistema com uma detecção baseada em padrões históricos de movimento, pois em alguns casos a cabeça de uma determinada pessoa não será localizada pela silhueta, já que pode estar no interior da silhueta de uma outra pessoa. Por fim, como o interesse é nas pessoas que estão olhando para os avisos, suas pupilas são detectadas por meio de iluminação infravermelha para que seja possível saber se a atenção está voltada ou não para o quadro de avisos. Os resultados obtidos são, então, comparados para que não haja contagem de uma pessoa mais de uma vez.

Lu [LU 01], Kettner [KETTNER 99] e Shapiro [SHAPIRO 01] localizam as pessoas através da **análise do comportamento dos pixels** seja ela efetuada através da diferença, do movimento ou da similaridade entre eles. Este tipo de análise também possibilita a localização / contagem de pessoas em grupos (i). Lu [LU 01] utiliza um processo que se baseia nos pixels que se modificaram do quadro anterior para o quadro atual. Primeiramente, o sistema estabelece um retângulo que envolve o maior número de pixels modificados possível. Em seguida, calcula um percentual para saber se a maior parte dos pixels deste

retângulo é relativa ao primeiro plano. Em caso positivo considera-se que uma pessoa foi encontrada, caso contrário o retângulo é dividido e observa-se novamente se a cobertura dos novos retângulos encontrados diz respeito principalmente a pixels do primeiro plano. Este procedimento se repete até que cada um dos retângulos encontrados contenha essencialmente pixels do primeiro plano de forma que cada um deles seja associado a uma pessoa. Caso uma determinada pessoa já tenha sido previamente localizada o novo retângulo definido para ela é eliminado e o antigo é mantido. Com este método, pode-se localizar pessoas em movimento, facilitar a construção do histograma relativo a cada uma delas e indicar possíveis oclusões.

Já Kettner [KETTNER 99] realiza a localização / contagem de pessoas através de um algoritmo adaptado de Boykov [BOYKOV 98] que segmenta a imagem em regiões máximas cujos pixels se assemelham e se movimentam de maneira similar aos pixels vizinhos. Este processo faz com que sombras claras e reflexos sejam tratados como fundo, e não como parte das pessoas.

Shapiro [SHAPIRO 01] se baseia na movimentação dos pixels entre dois quadros consecutivos. Para isto, precisa, inicialmente, identificar os pixels que se movimentaram de um quadro para outro. Esta identificação começa com a limiarização da diferença entre dois quadros consecutivos. Depois deste passo, os componentes conectados são extraídos e as regiões pequenas assumidas como ruídos são removidas. Em seguida, uma operação morfológica de fechamento é implementada e os retângulos mínimos envolventes (bounding boxes) das regiões remanescentes são encontrados. Com isto, os pixels em movimento serão aqueles encontrados na limiarização juntamente com os relativos aos retângulos envolventes mínimos.

A próxima etapa é definir alguns pontos estratégicos dentre os pixels em movimento, chamados pontos de interesse. A determinação destes pontos é feita através da análise de variância. Uma vez cumprida esta etapa, deve ser encontrado um ponto correspondente no quadro seguinte para cada ponto de interesse. Esta correspondência é feita através da correlação cruzada entre uma região de interesse, que considera a vizinhança do ponto de interesse, e uma região de busca um pouco mais ampla na imagem seguinte. Após o estabelecimento destas correspondências, vetores de movimento são encontrados para cada ponto estratégico, de modo que uma pessoa ou objeto em movimento é localizado no quadro seguinte.

Os sistemas de Hayfron [HAYFRON 01] e de Ramanan [RAMANAN 03] possuem técnicas distintas, mas têm em comum o fato de tratarem a distinção

de objetos e de não abordarem a aplicação do sistema em casos onde as pessoas já entram em cena em grupos (i). Hayfron [HAYFRON 01] faz a identificação de pessoas através da **simetria do movimento**, uma vez que a disposição dos membros do corpo humano nos leva a movimentos simétricos. Para tanto, encontra mapas representativos da simetria das pessoas a partir da aplicação de um operador de simetria na imagem obtida após a aplicação do operador de gradiente na etapa de separação do fundo / primeiro plano. Em seguida, calcula a média dos mapas de simetria de uma seqüência de imagens, tal média é chamada de assinatura e através da análise desta será possível constatar se o objeto em questão representa ou não uma pessoa.

Ramanan [RAMANAN 03] utiliza um **modelo gráfico probabilístico** para identificar a configuração de corpos nos primeiros quadros da seqüência de imagens. O método busca candidatos a segmentos do corpo e, em seguida, os agrupa de forma a encontrar configurações que representem o corpo de uma pessoa e pareçam similares com o passar do tempo. Grupos que tenham dinâmica improvável são excluídos. Após esta etapa de aprendizado de aparências, vem a fase de busca nos próximos quadros das pessoas referentes aos modelos aprendidos, o que se dá através de uma rede Bayesiana. O sistema parece ser bastante robusto por ter um modelo de pessoa bastante completo.

Apesar de Wren [WREN 97] fazer uso da **análise de silhueta** em busca de formas relativas a cabeças para estabelecer a presença de uma pessoa, ele não atende ao item i) citado, pois, assim como Rossi [ROSSI 94] e Atsushi [ATSUSHI 02], é um sistema para pessoas isoladas, não tratando, portanto, a análise de grupos. Rossi [ROSSI 94] utiliza a **análise dos níveis de cinza** para detectar os limites das pessoas nas regiões de interesse, de forma que o fundo não pode ter muita textura ou, ao menos localmente, deve ser menos variado que as pessoas. Já Atsushi [ATSUSHI 02] localiza novas pessoas através das possíveis **posições simuladas** onde uma nova pessoa possa aparecer. Na medida em que o quadro atual tiver um objeto que coincida com algum objeto dos quadros de posição simulados considera-se que uma nova pessoa entrou no campo de visão da câmera.

Uma vez que as pessoas foram localizadas, as trajetórias delas podem ser encontradas. Alguns trabalhos, tais como os encontrados em [ROSSI 94], [PIAU 02], [LU 01], [CAI 95], [KETTNAK 99], [HARITA 00], [WOJTASZ 02], [WREN 97], [ATSUSHI 02], [SEGEN 96], [HAYFRON 01], [RAMANAN 03], [ROH 00], [BEN 02], [COLLINS 02], [POON 02] e [KAMIJO 02] tratam este aspecto de diferentes

formas. Outros autores possuem ainda trabalhos que envolvem a localização de partes do corpo e não apenas do corpo inteiro ([POLAT 01], [FRIGOLA 03], [LEE 03], [LEE 02], [XIAO 03], [RAJARAM 03], [FORSYTH 03] e [PARK 00]).