

Navegação Robótica Aérea Baseada em Visão – Requisitos de Processamento de Imagens para Projeto e Implementação

JOSÉ LUIZ DE SOUZA PIO^{1,2}

¹VerLab – DCC – ICEx UFMG , Av. Antônio Carlos, 6657, Belo Horizonte, MG, Brasil, 31270-901

²DCC – FUA – Universidade do Amazonas, Av. Gen. Rodrigo Otávio J. Ramos, 3000, Manaus, AM, 69077-000
josepio@dcc.ufmg.br

Resumo: Este trabalho apresenta os conceitos básicos e o projeto de um sistema para a navegação de um robô aéreo do tipo dirigível. São configurados, em uma primeira análise, os algoritmos e técnicas envolvidos na operacionalização do dirigível robótico em ambiente natural e externo. É destacada também a teoria que envolve a implementação de sistemas para a obtenção da informação visual enfatizando-se os aspectos inerentes a extração de características e o processamento de imagens em tempo real.

1 Introdução

A navegação é a atividade básica de um robô móvel, habilitando-o ao livre deslocamento pelo seu espaço de trabalho, alcançando metas pré-definidas, realizando tarefas e desviando de obstáculos [Latombe (1991)]. Isto é obtido através de sistemas de navegação, presentes em todos os robôs deste tipo, algumas vezes diferentes quanto à forma de executar o seu trabalho, mas chegando objetivamente geração ou rastreamento de uma trajetória evitando o choque com obstáculos e preservando a integridade do robô.

Os sistemas de navegação autônoma baseados em visão consistem, basicamente, de um veículo robótico (aéreo, terrestre, aquático ou sub-aquático), dispositivos de captura de imagens (câmeras e *frame grabbers* para a aquisição e digitalização de imagens) [Kundur (1998) e Kundur (2000)]. O sistema é formado por dois subsistemas, um geralmente embarcado que é o responsável pela aquisição de imagens e extração das informações visuais importantes e outro subsistema (referenciado também como sistema de controle), que pode ser embarcado ou não dependendo das características do robô como sua capacidade de carga e energia, que gera os comandos apropriados para o veículo descrever uma navegação segura, em tempo real, baseada na informação visual obtida.

A principal característica dos sistemas de navegação baseados em visão é a necessidade de extrair informações visuais relevantes da cena imageada em tempo real. Estas informações são convertidas em instruções pelo sistema de controle que ativa ações específicas no robô. Estes sistemas funcionam de forma cíclica, provendo um fluxo de informações constante em malha fechada [Kundur (2000)]. O sistema deve ser capaz ainda de prover a posição e orientação do robô com base em imagens captadas por sua câmera.

O controle, por sua vez, terá de atuar sobre os propulsores do robô de modo a fazê-lo seguir trajetórias pré-programadas, baseado, evidentemente, nas informações visuais extraídas das imagens. O sistema de controle deve ser suficientemente hábil para desviar de obstáculos e corrigir a rota do veículo, necessitando de um processo que favoreça, a algum nível, a tomada de decisões. Essa tomada de decisões implica na necessidade de um modelo dinâmico do robô que dê ao sistema de controle meios de calcular e prever os comandos a serem enviados aos propulsores.

Uma classe importante de robôs são os chamados robôs aéreos, os quais deslocam-se no espaço tridimensional, em ambientes externos ou internos através de vôo com propulsão ou flutuação. Este trabalho apresenta uma visão geral sobre as técnicas empregadas em sistemas navegação de robôs aéreos baseados em visão, destacando os aspectos inerentes ao processamento de imagens sistematizando respostas a duas questões fundamentais: (i) Que informações visuais são importantes e devem ser extraídas de uma seqüência de imagens? (ii) Como se procede para extrair estas informações?

O artigo está estruturado em cinco seções. A primeira seção apresenta o problema que é tratado neste artigo e descreve em linhas gerais os sistemas de navegação baseados em visão. A fundamentação teórica é tratada na segunda seção. Uma breve descrição dos trabalhos recentes sobre sistemas de navegação baseada em visão para veículos aéreos é

tratado na terceira seção. A quarta seção descreve os requisitos de hardware e software para um sistema de navegação do dirigível robótico. A conclusão e algumas considerações para trabalhos futuros é realizada na quinta e última seção.

2 Fundamentos Teóricos

Sistemas servo visuais são formados pela utilização de sensores de visão em um sistema de controle de malha fechada. As abordagens usuais para esses sistemas são classificadas em duas diferentes categorias [Hutchinson(1996)]: baseadas em posição e baseadas em imagens.

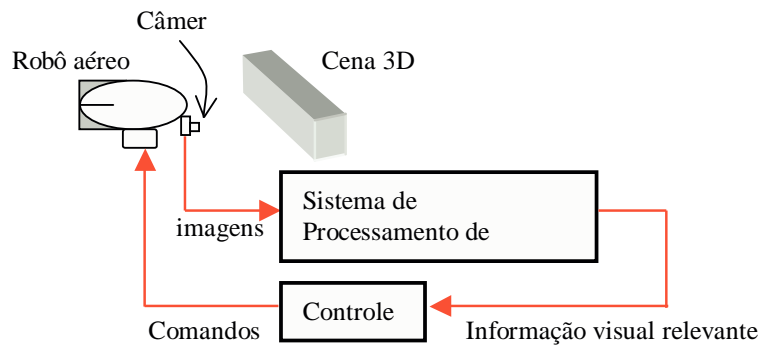


Figura 1: Sistema de controle servo visual

No controle baseado em posição as características são extraídas da imagem e usadas em conjunção com o modelo geométrico do alvo. Também faz-se necessário o conhecimento do modelo da câmera para a determinação do posicionamento do alvo em relação a esta. A retroalimentação é computada para a estimativa dos erros de posicionamento dos alvos no espaço. O controle servo visual baseado em imagem visa minimizar a complexidade computacional de tempo eliminando a necessidade de interpretação da imagem e os erros oriundos do modelo e calibração da câmera. Conceitos básicos relacionados a forma de aquisição e processamento da informação visual desses sistemas são descritos nas seções seguintes.

2.1 Modelos de Câmeras e Projeções

Para se controlar um robô usando as informações providas por um sistema de visão computacional, é necessário a compreensão dos aspectos geométricos do processo de formação da imagem. Os modelos adotados para câmeras são os de furo (*pinhole camera*), sendo estas compostas por um conjunto de lentes que formam a projeção 2D da cena imageada. O processo de projeção pode ser a projeção perspectiva ou ortográfica (Figura 2).

O processo de projeção causa a perda da informação de profundidade, uma vez que cada ponto projetado no plano de imagem corresponde a uma intensidade de brilho dos pontos tridimensionais que compõe os objetos que formam a cena. Logo, a obtenção da informação de profundidade só pode ser alcançada se for admitida alguma informação adicional. Esta informação pode ser obtida com a utilização de múltiplas câmeras, múltiplas visadas com uma única câmera ou através do conhecimento das relações geométricas entre cada ponto na imagem e os objetos que compõe a cena.

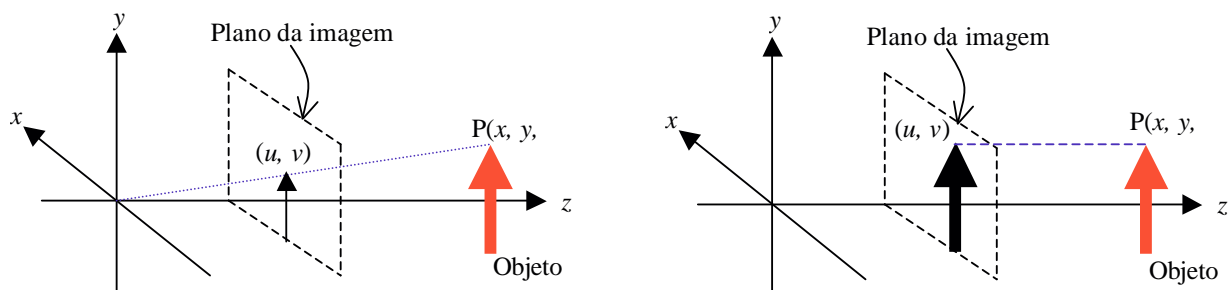


Figura 2: Projeção perspectiva e projeção ortográfica

Câmeras podem ser modeladas como transformações de sistemas de coordenadas tridimensionais. Essas transformações incluem componentes de translação (especificada pelo centro de projeção, usualmente o centro óptico da câmera), rotação (ângulos de *pan*, *tilt*, *roll*), escala e perspectiva. O estabelecimento dos parâmetros extrínsecos (rotação e translação) e intrínsecos (perspectiva e escala) da câmera é chamado calibração. Esses parâmetros podem ser calculados a partir do conhecimento de coordenadas 3D de alguns pontos de referência (*test-field calibration*) ou pela combinação de alguns pares de pontos característicos nas duas imagens (*self calibration*). Vários métodos de calibração estão registrados na literatura, assim como maneiras de se eliminar essa etapa na obtenção de informações de objetos através de sua imagem.

2.2 Características da Imagem e Espaço de Parâmetros de Características

Uma característica de uma imagem é qualquer característica estrutural que pode ser extraída de uma imagem como uma aresta ou quina. Tipicamente características de imagens correspondem a projeções das características físicas dos objetos no plano de imagem da câmera [Hutchinson (1996)]. Um parâmetro característico de uma imagem é qualquer valor real calculado a partir de uma ou mais imagens. Estes valores incluem áreas, momentos, relacionamentos entre regiões ou vértices, etc. Boas características são aquelas que podem ser localizadas de forma não ambígua sob diferentes vistas de uma cena.

2.3 Extração da Informação Visual

Quando um sistema de navegação autônoma identifica as informações visuais contidas em uma imagem, a extração dessas informações passa a ser uma tarefa importante. As abordagens para a extração da informação visual são classificadas por [Kundur (1998)] como baseadas em reconstrução e abordagens não baseadas em reconstrução.

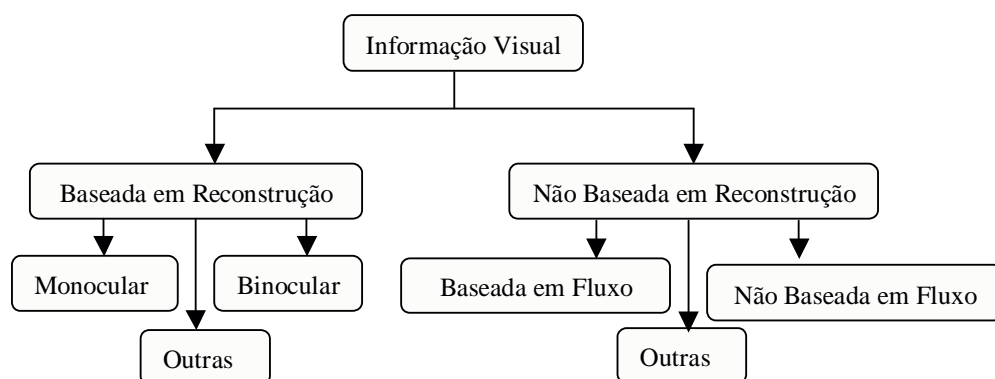


Figura 3: Abordagens de extração da informação visual

A Figura 3 apresenta um esquema através do qual se visualiza as abordagens para a extração da informação visual. As técnicas de reconstrução são decompostas de acordo com o número de câmeras utilizadas (uma ou mais câmeras). Por sua vez, o conjunto de técnicas que não são baseadas em reconstrução utilizam informações obtidas da sensação de movimento (fluxo) ou nas características identificáveis nas imagens, como pontos, segmentos de retas, figuras geométricas, etc.

Os métodos baseados em reconstrução obtêm características da imagem e recompõe a cena através da construção e dos objetos que formam a cena. Técnicas dessa natureza são mais robustos e convenientemente empregados nos sistemas de visão.

2.4 Abordagens Baseadas em Reconstrução para a Navegação Autônoma

A seguir são sumariamente apresentadas as principais técnicas de reconstrução para a navegação autônoma.

Estéreo

A análise de imagens em estéreo constitui num importante método passivo para extrair a estrutura tridimensional de uma cena. O número de câmeras usadas (duas ou mais) determina a geometria da imagem – suas posições e orientações – e é fator importante de determinação da disparidade de itens correspondentes em imagens estéreo, e também para o cálculo das linhas epipolares, como descrito a seguir.

Quanto à orientação, as câmeras podem estar arrançadas de forma convergente ou com seus eixos paralelos. A geometria de imageamento estéreo convencional envolve um par de câmeras com eixos ópticos paralelos e separadas por uma distância horizontal chamada *linha de base estéreo* (*stereo baseline*). Os eixos ópticos são perpendiculares à linha de base e as linhas da imagem (*scanlines*) são paralelas à linha de base. Caso as câmeras não estejam arrançadas segundo esta geometria padrão, é possível conseguir tal configuração através da aplicação de transformações afins sobre as imagens, etapa denominada *retificação*. De acordo com a escolha das características obtidas no par de imagens, pode-se aplicar dois tipos de algoritmos: Baseados em áreas e baseados em características.

Os algoritmos baseados em área partem de um ponto de interesse escolhido em uma imagem, aplica-se, então, uma medida de correlação para procurar por ponto correspondente com uma vizinhança que combine na outra imagem pelo uso direto de medidas de intensidade. A desvantagem é ser sensível a variações fotométricas e distorções oriundas da mudança do ponto de vista.

Por sua vez, os algoritmos baseados em Características utilizam pontos de borda, segmentos de reta, curvas, contornos, regiões e estruturas topológicas que são escolhidas e combinadas. Sua vantagem é ser menos sensível a variações fotométricas, ser mais rápido devido ao menor número de candidatos e a possibilidade das características serem extraídas com precisão de sub-pixel. A desvantagem é o pequeno número de estruturas correspondentes devido ao pequeno número de características.

Com base na visão estereoscópica, podem ser enumerados os seguintes componentes para inferência de informação tridimensional de imagens: a) aquisição de imagem e modelagem de câmera – determinar o nível de resolução e precisão, e os parâmetros da geometria do imageamento suficientes para embasar o algoritmo de combinação; b) seleção de características – determinar os elementos primitivos a serem usadas como base para o algoritmos de combinação; c) correspondência – combinar os elementos nas duas imagens que são projeções do mesmo elemento no mundo 3D; d) interpretação da disparidade – converter a disparidade em informação de profundidade. Esta técnica é bastante empregada na navegação robótica, principalmente na construção de mapas e modelos digitais de terreno, como no trabalho de Lacroix(2002).

Foco

Mudanças no ajuste do foco da lente resultam numa variação do grau de nitidez da imagem. Minimizando-se o borramento da imagem, a profundidade da superfície pode ser estimada. O mecanismo de foco ajusta a distância entre o centro da lente e o plano de imagem da câmera. Assim, a distância u entre o objeto e o centro da lente de distância focal f e a distância v entre o lado oposto da lente e o plano de imagem podem ser relacionadas pela equação $1/u + 1/v = 1/f$.

Assim, se um ponto na imagem está em foco e se v e f são conhecidos, a partir desta equação (chamada equação da lente) pode-se determinar a distância ao objeto. Quando um ponto da cena não satisfaz a equação da lente, a imagem está borrada. O efeito da imagem desfocada pode ser modelada, em uma primeira aproximação, pela convolução da imagem com um filtro passa-baixa, o que implica na perda de frequências espaciais altas da imagem. Medições de altas frequências podem portanto ser usadas para desenvolver uma função objetiva que assuma seu valor ótimo quando o borramento da imagem é minimizado.

No entanto, a precisão do método é limitada pelos fenômenos *chamados profundidade de foco e profundidade de campo*. Um ponto pode borrar em um pequeno círculo antes mesmo que a perda de nitidez seja detectada. Também, devido à capacidade limitada do sistema em discernir elementos da imagem, o plano de imagem pode se mover por um intervalo de localizações próximo a v (ou mesmo para uma mesma localização v na imagem, objetos podem se mover em um certo intervalo de profundidades – chamado profundidade de campo) e a imagem parecer igualmente em foco.

Textura

Da mesma forma que a intensidade em um pixel varia com a mudança de orientação da superfície, varia também a aparência, isto é, a forma e densidade local de um *texel* (elemento de textura). A reconstrução a partir da textura também precisa de algumas suposições. Uma dessas suposições é que o padrão original da superfície é homogêneo e qualquer variação observada na imagem, como distorção da forma e densidade da textura, é devida à mudança da forma da

superfície e do ponto de vista (restrição da homogeneidade da textura). Essa suposição restringe a orientação das regiões mas não é suficiente para determiná-las de forma única, sendo portanto necessárias suposições geométricas para propagar as restrições e facilitar a busca por uma solução globalmente correta. Um método de se determinar a orientação de uma superfície a partir de sua textura de linhas paralelas, considerando que linhas convergentes em uma imagem sob projeção perspectiva são originalmente linhas paralelas em uma única superfície planar. Mesmo se elas não se encontram aí explicitamente, linhas paralelas podem ser também obtidas de uma textura se forem feitas restrições sobre a forma ou a distribuição dos elementos dessa textura. Se for tomada, por exemplo, uma textura com padrões circulares de mesmo tamanho, ou mesmo uma textura com vários tipos de padrão mas que sejam relativamente homogêneos e possuam mesma orientação (uma parede de tijolos por exemplo), pode-se determinar a orientação da superfície considerando-se as retas que passam pelo centro de massa dos elementos do padrão e os pontos de fuga correspondentes. No entanto este método não é eficiente se a orientação dos elementos da textura não for uniforme, o que restringe sua aplicação. Alguns trabalhos, como o de Lecroix (2002) por exemplo, utilizam a textura para identificar regiões homogêneas e de possível correspondência entre pares de imagens estéreo.

Tonalidade

A técnica de forma a partir da tonalidade (*shape from shading*) consiste na recuperação da forma da superfície a partir de uma única imagem e do conhecimento do mapa de reflectância de uma determinada região da superfície [Horn (1985)]. Em geral, a intensidade de cada pixel na imagem depende do material que compõe o objeto imageado, posição e orientação da superfície, posição da iluminação e posição do observador. Se, no entanto, se considerar que a iluminação incidente em cada posição da superfície é paralela e constante, e que a projeção da imagem é ortográfica, as características de reflectância da superfície de um objeto podem ser representadas como $f(i, e, g)$, isto é, em função dos ângulos entre a luz incidente e a normal à superfície (i), o ângulo entre a luz emergente (que é também a linha de visada) e a normal à superfície (e), e o ângulo de fase (ângulo entre a luz incidente e a luz emergente – g). O mapa de reflectância é a função que relaciona a radiância da região (potência luminosa emitida por unidade de área, por unidade de ângulo sólido), com sua orientação local, a iluminação da cena e a posição do observador. Um dos modelos mais simples de reflectância é a perfeitamente difusa (Lambertiana) na qual a reflexão é uniforme para todos os ângulos de visada, e a quantidade de reflexão varia com o cosseno do ângulo incidente, isto é: $I(x, y) = f(i, e, g) = R(p, q) = \rho \cos i$, onde $I(x, y)$ é a intensidade do pixel de coordenadas da imagem (x, y) ; $R(p, q)$ é o valor do mapa de reflectância para a região com orientação (p, q) no espaço gradiente; ρ é a constante de refletividade. Uma das formas de se obter um mapa de reflectância é determiná-lo empiricamente, ou seja, utilizando-se um objeto do mesmo material daquele cujo mapa de reflectância se deseja obter, mas cuja forma seja conhecida, e registrando-se os valores de intensidades resultantes sob uma dada iluminação para cada orientação no objeto. Os objetos escolhidos para a calibração, como é conhecida esta técnica, devem ter na sua superfície visível pontos que compreendam uma grande faixa de orientações (p, q) , como é o caso da esfera. O *shape from shading* é de difícil aplicação para cenas genéricas, compostas por objetos com diferentes tipos de superfícies e com fontes de luz de orientação desconhecida. Seu uso, portanto, se limita a configurações de ambientes fechados, com iluminação controlada, sendo aplicado na navegação robótica somente os conceitos de reflectância em conjunto com outras técnicas, como as técnicas de reconstrução a partir do foco, por exemplo.

2.5 Abordagens Não Baseadas em Reconstrução

As abordagens para a navegação não baseadas em reconstrução não necessitam da estrutura explícita da cena e podem ser classificadas, segundo Kundur (1998), em abordagens baseadas em fluxo e em abordagens não baseada em fluxo. Por fluxo entende-se as variações sensoriais induzidas na retina de um observador pelo movimento relativo entre tal observador e a cena 3D. Este padrão de variação na retina do observador é usualmente referenciado como padrão de fluxo, fluxo ótico ou fluxo de imagem.

A idéia geral dessas abordagens é a de obter a informação de profundidade através da análise múltiplas vistas da cena. As variações entre essas imagens, decorrentes do movimento relativo entre câmera e objetos que compõem a cena, são usadas como formas para recuperação da profundidade, e na sua forma mais simples utiliza o movimento entre o observador e o ambiente tridimensional, que envolve os seguintes parâmetros: três velocidades translacionais v_x , v_y e v_z e três velocidades angulares w_x , w_y , w_z . O padrão de fluxo correspondente na retina do observador pode ser descrita por dois parâmetros u_x , u_y .

A utilização deste tipo de abordagem têm sido bastante empregada em navegação robótica no desvio de obstáculos, isto é obtido através da observação dos parâmetros u_x e u_y quando o robô muda sua posição frente a possível colisão com algum obstáculo. O problema geral no contexto da navegação é a determinação, ou a extração dos valores de u_x e u_y de uma sequência de imagens.

As abordagens não baseadas em fluxo incluem a extração de propriedades geométricas como área, forma, tamanho, contorno, etc; propriedades texturais como a densidade, contraste, etc; propriedades da câmera como foco, resolução, zoom, etc.

3 Trabalhos Relacionados

Coelho (1998) propõe um Sistema de Visão que tem por objetivo prover em tempo real a posição e orientação do dirigível com base em imagens captadas por sua câmera. Para isso, ele se vale do conceito de marco visual, um objeto com propriedades geométricas previamente conhecidas. A partir da imagem de um objeto são calculados através de suas projeções as suas posição e orientação. O marco visual é formado por três círculos com distâncias conhecidas entre os centros. Cada círculo é rotulado com um número de buracos variando de zero a dois. Os círculos são pretos e o marco é colocado contra um fundo claro a fim de facilitar a segmentação de imagem. Esta forma geométrica facilita a definição de um sistema de coordenadas do mundo, a partir do qual o robô pode se orientar em seu espaço de trabalho.

Em trabalho mais recente, Altug (2002) apresenta o controle de um helicóptero quadrotor (com quatro rotores) usando retroação visual. Este sistema utiliza uma câmera para estimar a posição e a orientação do helicóptero. O sistema de controle é externo (não embarcado) que recebe as imagens da câmera processando as informações e enviando os comandos para o processador embarcado do helicóptero, via *link* de rádio. O processador embarcado, por sua vez, tem objetivo checar e executar os comandos enviados confrontando-os com as informações do giroscópio, ajustando a velocidade de acordo com a posição e orientação desejada. O sistema utiliza marcos visuais, figuras circulares coloridas com 2,5 cm de raio, posicionadas em baixo do quadrotor. Um algoritmo de rastreamento é utilizado para encontrar a posição e calcular as áreas das figuras no plano da imagem.

Lacroix (2002) apresenta uma abordagem cooperativa que utiliza robôs aéreos e terrestres para a navegação e construção de mapas digitais de elevação. Um robô aéreo determina as posições relativas entre duas imagens estéreo sucessivas de um robô terrestre. Um algoritmo de recuperação do movimento é desenvolvido para rastrear as diferentes posições do robô terrestres, e através da integração dos dados obtidos pelo dirigível e pelo robô terrestre se obtém o mapa digital de elevação.

A descrição de um sistema servo visual para a decolagem de um dirigível é apresentada em Azinheira (2002). Neste trabalho o destaque é dado para a modelagem dinâmica do robô. Os aspectos visuais considerados utilizam os marcos visuais que favorecem, através das distorções providas pela projeção desses marcos no plano da imagem, o processo de decolagem do dirigível robótico.

A coordenação de operações de robôs terrestres com o auxílio de um robô aéreo é o objetivo de Kim (2002). Neste trabalho é descrito um sistema de gerenciamento de voo embarcado em um helicóptero robótico que propicia a realização de rastreamento de veículos terrestres. Um grupo de veículos terrestres é coordenado através de comandos que são passados pelo robô aéreo. Assim, um possível robô em fuga é rastreado e a estratégia de busca é passada para os robôs de terra pelo veículo aéreo dotado de um sistema de visão.

O desenvolvimento de sistemas de navegação e controle de dirigíveis é um tópico bastante explorado pelo grupo de pesquisadores que compõe o projeto AURORA. Referências interessantes sobre o controle e aplicação deste projeto são explorados em Maeta (2001), Pereira (2001) e Ramos (2000). Os aspectos relacionados com o sistema servo visual são tratados em Azinheira (2002).

4 O Projeto do Sistema de Navegação Visual para um Dirigível Robótico

Os robôs aéreos ou veículos aéreos não tripulados (do inglês, *Unmanned Aerial Vehicles* -- UAV), representam uma classe de sistemas robóticos móveis que combinem a capacidade de voo com o poder de processamento, dotado de autonomia para decidir rotas de navegação e estratégias para a realização de tarefas pré-determinadas. Os dirigíveis robóticos representam uma classe específica de UAV's que possuem a capacidade flutuar e se deslocar no espaço tridimensional, sendo idealmente empregados para a realização de tarefas que requeiram maior estabilidade, silêncio e pouca vibração.

O dirigível robótico deve ser capaz de navegar autonomamente através da obtenção das trajetórias impostas pela tarefa que deve ser executada. As trajetórias podem ser previamente definidas ou definidas em tempo real, no decorrer da navegação. Porém, em ambos os casos, algumas considerações devem ser observadas, como o estado interno do dirigível, e as condições do ambiente, como a direção dos ventos, relevo, e o posicionamento de obstáculos.

Á maneira convencional de se definir previamente a trajetória do dirigível é estabelecendo as coordenadas de chegada (objetivo) associadas a uma altitude de voo, geralmente levando-se em consideração o relevo da região, e decompondo o voo em decolagem/aterissagem, cruzeiro ou pairado.

Quando a navegação dá-se em tempo real, ou seja, as informações são usadas pelo sistema de navegação para definir uma trajetória de forma exploratória são obtidas a partir de características locais imageadas e detectáveis pelo sistema, como estradas, rios, tubulações, linhas de transmissão de energia, ou marcas identificáveis no terreno.

Um sistema de controle tipo servo visual deve possibilitar a geração de instruções para o controle de movimento e métricas para o controle de voo, como a altitude e posição do veículo frente a um sistema referencial global. O sistema funciona em uma plataforma externa, estando o dirigível equipado com uma câmera e um *link* de rádio para a plataforma externa (Figura 4). A câmera do dirigível captura as imagens que são transmitidas para a plataforma externa que, por sua vez, processa as imagens e transforma a informação visual em instruções que são enviadas para o processador embarcado do robô.

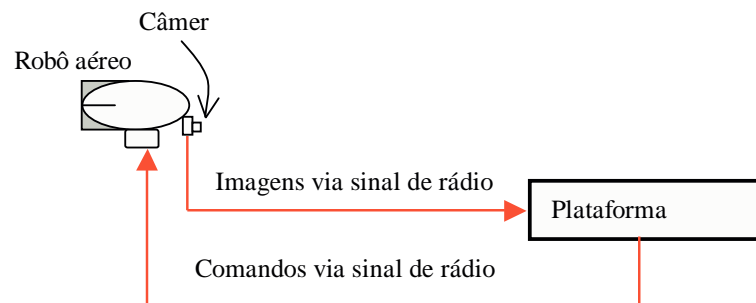


Figura 4: A plataforma externa é responsável pelo processamento principal

4.1 Metodologia para o Processamento de Imagens

Como metodologia geral para o processamento de imagens do sistema de navegação, optou-se em dividir o processo em três etapas: 1. Extração de características; 2. Matching; 3. Posicionamento e direção.

A Figura 5 apresenta o fluxo de informações obtido através do processamento realizado em cada etapa. A partir da imagem capturada pela câmera as características são extraídas e enviadas para o processamento de correspondência, através do qual gera as informações que indicarão o posicionamento e a direção do robô.

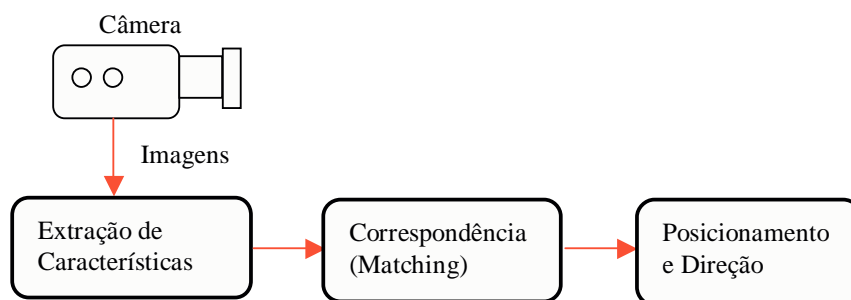


Figura 5: Etapas de processamento das imagens do sistema de navegação baseado em visão

Assim, cada etapa de processamento de imagens fica com uma função específica (Figura 5). A primeira etapa prevê a identificação de características detectáveis no terreno, segmentos de retas no primeiro instante. Diferentes algoritmos são aplicáveis a esta etapa que visa a segmentação da imagem em diferentes segmentos de reta. Uma abordagem usual é, por exemplo, a identificação de regiões de linhas de suporte. Nesta abordagem, pixels que possuem valor do gradiente maior que determinado limiar são agrupados em regiões com direções similares de gradiente. Uma superfície planar de brilho é aproximada, por mínimos quadrados, a cada região previamente determinada, ajustando o valor do brilho em

função das coordenadas da imagem. Outra técnica seria o emprego da transformada de Hough para a identificação dos segmentos de reta.

Uma outra característica importante é a detecção de pontos, normalmente relacionados a arestas dos objetos poligonais que possam eventualmente compor a cena de um ambiente interno ou ponto de cruzamentos de vias em ambientes urbanos externos. A primeira idéia é a de se obter os pontos a partir dos valores dos limites dos segmentos de retas previamente obtidos. Porém, a idéia geral é a identificação de dos pontos com curvatura máxima em uma cadeia de arestas.

A etapa de correspondência (*matching*) prevê a movimentação do dirigível a partir da mudança da posição das características anteriormente detectáveis na etapa anterior (segmentos de retas e pontos). Deve-se para tanto rastrear estas características através de um loop de predição-correspondência-atualização. A captura da movimentação e da estrutura da cena é representada através de um processo natural de projeção, ou seja, assume-se que o modelo de câmera é do tipo *pinhole* e que a origem do sistema de referência da câmera é centro de projeção; o eixo *z* está alinhado com o eixo focal. A localização de uma característica na imagem é obtida através dos parâmetros de calibração da câmera.

A movimentação pode ser efetivada através da identificação de pontos em segmentos de reta. Grupamentos de retas podem identificar, devido a projeção perspectiva, pontos de fuga que indiquem a direção e o sentido da navegação. Um exemplo típico disto é a interseção de duas retas paralelas, como as laterais de uma estrada, que se encontram no horizonte. O ponto de fuga definido no horizonte indica a direção da navegação.

Atualmente, algumas amostras de imagens estão sendo tomadas para a implementação desses algoritmos em ambientes naturais, objetivando a adequação e aprimoramento das técnicas aqui apresentadas. O VerLab (Laboratório de Visão Computacional e Robótica) da UFMG possui um dirigível robótico que favorece a viabilização de experimentos e aprimoramento das técnicas a serem aplicadas no contexto da cooperação entre grupos de robôs.

5 Conclusão

Este trabalho apresentou algumas técnicas para a obtenção da informação visual necessária para o desenvolvimento de sistemas de navegação de robôs aéreos baseado em visão. Destacou-se os aspectos inerentes ao processamento de imagens e sua possível adequação a um robô dirigível equipado com uma única câmera e com *link* de rádio para uma plataforma externa, não embarcada, responsável pelo processamento das imagens e envio dos comandos de navegação ao robô.

O desenvolvimento bem sucedido do sistema aqui apresentado é uma prova conceitual da aplicabilidade dos dirigíveis autônomos em ambientes externos, naturais e extensos. Além disso, o sistema representa uma oportunidade de trabalho original multidisciplinar, envolvendo as áreas de Visão Computacional, Robótica Móvel e Processamento Digital de Imagens. Por outro lado, o esquema de navegação baseado em visão é extensível a outros tipos de robôs aéreos, como aviões e helicópteros, a robôs móveis terrestres e mesmo a robôs subaquáticos que possam, por exemplo, inspecionar um cabo ou duto submarino, desde que existam condições de visibilidade.

Entretanto, este trabalho é uma primeira abordagem ao um problema geral que visa integrar, através de sistema de visão, grupos de robôs aéreos e terrestres para a navegação e tarefas cooperativas. Como trabalhos futuros prevê-se a implementação da técnica aqui descrita, sua avaliação frente a situações reais (imagens de ambientes naturais com presença de ruídos) e a confrontação dos resultados com outras técnicas usuais, assim como a possível adequação de outras características como textura e fluxo para a viabilização da navegação.

6 Referências

- [1] ALTUG, E. "Control of a Quadrotor Helicopter Using Visual Feedback". ". *Proceedings of the 2002 IEEE International Conference on Robotics & Automation*, Washington, USA, maio, 2002.
- [2] AZINHEIRA, José R. et al. "Visual Servo for the Hovering of na Outdoor Robotic Airship". *Proceedings of the 2002 IEEE International Conference on Robotics & Automation*, Washington, USA, maio, 2002.
- [3] COELHO, L. S. E CAMPOS, M.F. M. "Navegação de Dirigíveis Autônomos Baseados em Visão Computacional". Obtido na WWW na URL: <http://www.dcc.ufmg.br/pos/html/spg98/anais/omni/omni.html>, 1998.
- [4] HORN, B., "Robot Vision", MIT Press, 1985.
- [5] HUTCHINSON, S. "A Tutorial on Visual Servo Control". Universidade de Illinois, 1996.
- [6] KIM, JIM H. "FLYING ROBOTS: Modeling, Control and Decision Making". ICRA, 2002.
- [7] KUNDUR, S. e RAVIV, Daniel, "A Vision-Based Pragmatic Strategy for Autonomous Navigation". *Pattern Recognition*, Elsevier Science, Vol. 31, N. 9, pg. 1221-1239, 1998.

- [8] KUNDUR, S. e RAVIV, Daniel, "Active Vision-Based Control Schemes for Autonomous Navigation Tasks". *Pattern Recognition*, Elsevier Science, N. 33, pg. 295-308, 2000.
- [9] LACROIX, S. ET AL. "Towards Cooperative Air/Ground Robotics: Issues Related to Digital Elevation Map Building". LAAS/CNRS, França, 2002.
- [10] LATOMBE, J.C. "Robot Motion Planning". Kluwer Academic Publishers, Boston, 1991.
- [11] MAETA, S.M. "Desenvolvimento de Infraestrutura Embarcada do Projeto AURORA". Dissertação MSc. UNICAMP, Julho, 2001.
- [12] ORTEL, C. "Machine Vision-based Sensing for Helicopter Flight Control". *Robotica*, Vol. 18, pp. 299-303, 2000.
- [13] PEREIRA, C.S. ET AL "Formulação de um Sistema de Navegação Inercial para um Dirigível Não-tripulado: Primeiros Resultados". XVI Congresso Brasileiro de Engenharia Mecânica – COBEM, 2001.
- [14] RAMOS, J.J.G. ET AL. "Project AURORA: A Status Report". 3rd International Airship Convention and Exhibition, 2000.
- [15] SAGÜÉS, C. "Motion and Structure for Vision-Based Navigation". *Robotica*, Vol. 17, pp. 355-364, 1999.
- [16] ZWAAN, S. "Visual Station Keeping for Floating Robots in Unstructured Environments", *Robotics and Autonomous Systems*, Vol. 973, pp. 1-11, 2002.