

Universidade de São Paulo
Instituto de Ciências Matemáticas e de Computação

SLAM Visual

Fernando Zuher Mohamad Said Cavalcante

Orientadora: Prof^a. Dr^a. Roseli Romero

Mestrado em Computação, ICMC-USP

Divisão da apresentação

1. Introdução

- O que é SLAM?
- Classificação adotada

2. Principais abordagens

- Baseado em primitivas geométricas
- Diretamente

3. Conclusões

4. Sugestões de melhoria

Divisão da apresentação

1. Introdução

- O que é SLAM?
- Classificação adotada

O que é SLAM?

- SLAM = *Simultaneous Localization And Mapping*
- SLAM Visual
 - Em essência SLAM Visual é um problema de estimação paramétrica de tempo real onde as observações são provenientes de câmeras de vídeo, a qual consiste em obter um conjunto de parâmetros que melhor explica um *fenômeno*, dado um modelo e uma sequência de imagens.

O que é SLAM?

- Vantagens do SLAM Visual
 - Não é necessário contato com o ambiente para sua medição, a abstração dos dados tem um amplo campo de cobertura, há grande quantidade de informação, etc.
- Duas abordagens principais:
 - Por meio de blocos de imagens
 - Cada imagem individualmente

O que é SLAM?

- Na estimação paramétrica geralmente são considerados 3 passos:
 1. Definir um modelo apropriado;
 2. Resolver o problema intrínseco de associação de dados relacionado;
 3. Resolver o sistema de equações não lineares associado.

O que é SLAM?

- Classificação na literatura:
 - Dada pelo tipo de filtragem empregada: Filtro de Kalman estendido, filtro de partículas, etc.
- Classificação adotada
 - Distinção das abordagens de acordo com os trabalhos que utilizam etapas intermediárias dos que não**
 - Assim...

Classificação adotada

- Baseado em primitivas geométricas
 - Utilização de primitivas geométricas como pontos, retas, contornos, etc.
- Diretamente
 - A abstração dos dados provém dos níveis de intensidade dos pixels da imagem.

Divisão da apresentação

2. Principais abordagens

- Baseado em primitivas geométricas
- Diretamente

Baseados em primitivas geométricas

- O processo de obtenção dos parâmetros pode ser dividido em três etapas sequenciais principais:
 1. extração de dados
 2. associação de dados
 3. estimação paramétrica

*O problema de associação de dados é separado do problema de estimação paramétrica.

Baseados em primitivas geométricas

1. extração de dados
 - O detector de primitivas ideal...
 - A avaliação de desempenho de um detector em particular é efetuada a partir de duas métricas principais: **precisão e repetibilidade**.
 - Dado que invariância total não existe, detectores em geral são baseados em limiares para decidir se uma dada primitiva está presente ou não na imagem.

Baseados em primitivas geométricas

1. extração de dados
 - Detector de Harris:



Baseados em primitivas geométricas

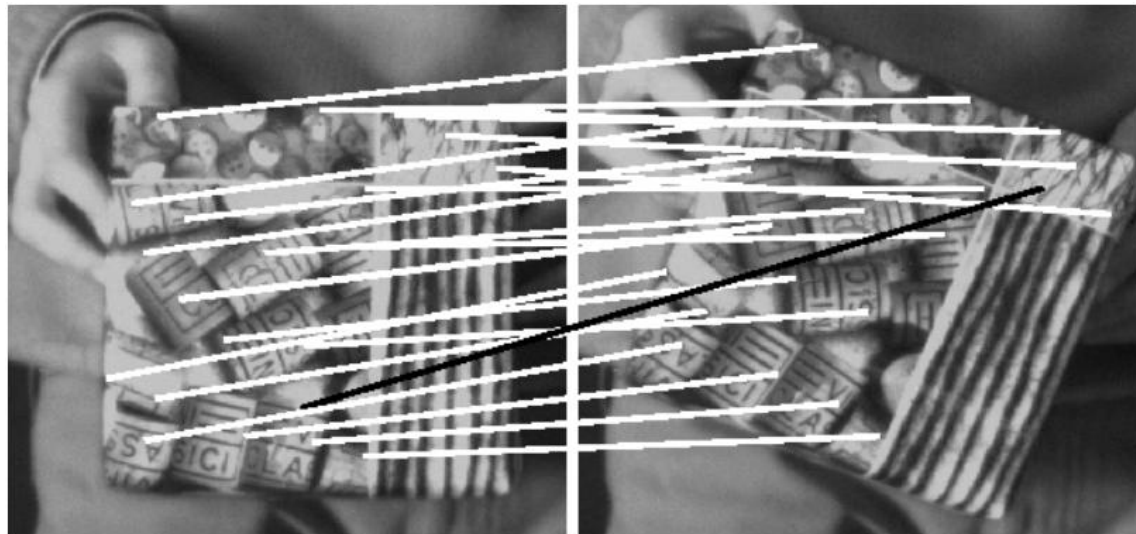
2. associação de dados

- A segunda etapa pode ser inicializada assumindo a hipótese de que um conjunto suficientemente grande das mesmas primitivas geométricas foram detectadas em imagens diferentes.
- Geralmente, a associação de dados é efetuada avaliando-se uma medida de similaridade entre descritores de cada uma das primitivas. Os descritores são em geral determinados utilizando uma região em torno da primitiva.

Baseados em primitivas geométricas

2. associação de dados

- Exemplo:



Baseados em primitivas geométricas

3. Resolver o sistema de equações não lineares associado.
 - Se existir um conjunto suficientemente grande de primitivas geométricas em correspondência nas imagens...
 - Cuidado especial com *outliers*.
- *Exemplo na figura anterior.

Baseados em primitivas geométricas

- Vantagens dessa técnica
 - Um ponto atrativo dessas abordagens se deve à divisão de um problema complexo em sub-tarefas**
 - Talvez isso explique sua vasta disseminação na literatura, a despeito de suas inerentes dificuldades.
 - A grande maioria das técnicas existentes pertencem a essa classe de abordagem.

Baseados em primitivas geométricas

- **Desvantagens** dessa técnica
 - Falsas correspondências, ajuste de diversos parâmetros, dependência de primitivas particulares, sensibilidade a mudanças fotogeométricas (como luminosidade), etc.

Principais abordagens

Diretamente

Diretamente

- O termo “direto” advém do fato que a intensidade do pixel é diretamente explorada para estimar os parâmetros relacionados.
- Por essa razão, eles são também chamados de técnicas baseadas em textura, baseadas na aparência, na intensidade, ou ainda em *templates*.

Diretamente

- Resolução simultânea do problema de associação de dados e o da estimação paramétrica.
- Em sua base, a abordagem direta ao SLAM Visual é realizada via registro direto de imagens.

Diretamente

- Exemplo:



Diretamente

- Seus valores ótimos são encontrados quando existir uma similaridade entre as intensidades dos pixels da imagem de referência e as da imagem corrente transformada.

Diretamente

- Nas técnicas diretas, para cada nova imagem capturada, é possível identificar três passos iterativos principais para resolver:
 1. Transformação da imagem corrente, dado um modelo e os parâmetros correntes;
 2. Obtenção da medida de similaridade, ou seja, o valor de uma função custo;
 3. Determinação do incremento nos parâmetros para diminuir o valor da função custo.

Diretamente

- Como a imagem inteira pode ser explorada o fator limitante para processá-la inteiramente recai sobre os recursos computacionais disponíveis.
- Se esses recursos forem poucos, é sempre possível efetuar uma escolha de quais regiões da imagem serão efetivamente exploradas.

Diretamente

- Vantagens
 - Em todo caso, a possibilidade de explorar toda a informação contida nas imagens (mesmo onde não há primitivas geométricas) constitui uma das vantagens mais importantes desta classe de métodos.

Diretamente

- **Desvantagens**
 - Por outro lado, métodos de otimização não linear que constituem a base das técnicas diretas de tempo real requerem uma inicialização adequada.
 - Na prática, isso significa que os deslocamentos dos objetos na imagem devem ser relativamente pequenos para evitar erros na solução (mais formalmente, cair em um mínimo local da função custo) ou, similarmente, que taxas mais altas de aquisição das imagens são necessárias.

Diretamente

- **Desvantagens**
 - Este último caso requer uma capacidade computacional maior para atender à essas altas taxas.
 - Isso constitui uma das maiores limitações dessa classe de métodos de estimação paramétrica.
 - Daí a necessidade de tratar adequadamente a inicialização do processo de otimização.

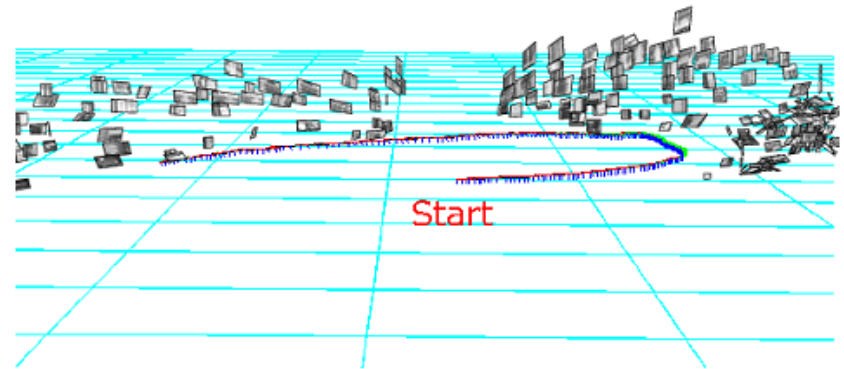
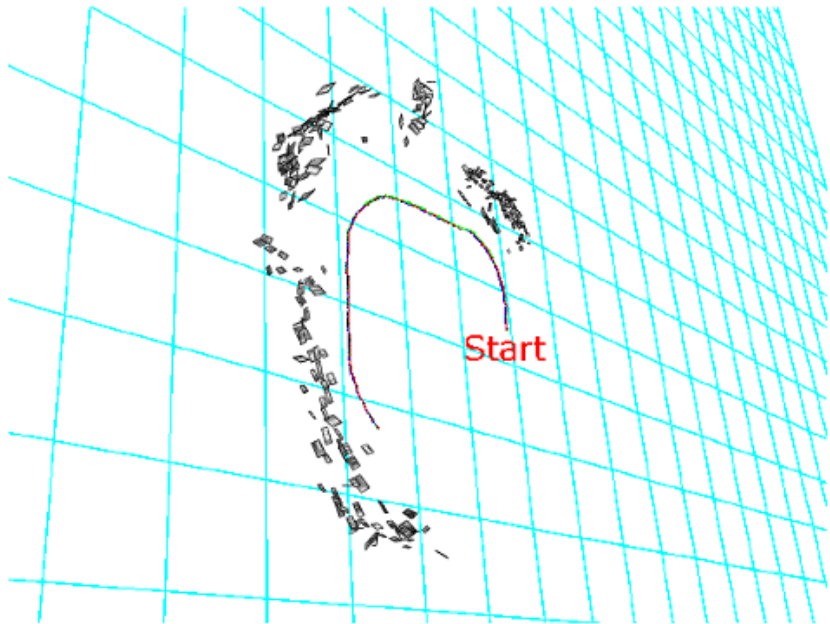
Diretamente

- Exemplo:



Diretamente

- Exemplo:



Divisão da apresentação

3. Conclusões

Conclusões

- O problema de SLAM monocular pode ser definido como um problema de estimação simultânea, e em tempo real, da pose da câmera e da estrutura da cena.
- Não obstante, o autor procura apresentar as duas principais classes de abordagens ao problema – as baseadas em primitivas geométricas e as ditas diretas –, e munir o leitor com “diversas referências”** que compõem o estado-da-arte.

Conclusões

- Em relação às tendências de pesquisa na área, diversas direções podem ser enumeradas:
 - Do ponto de vista de estimação paramétrica, existe bastante espaço para o estudo de quais técnicas são as mais apropriadas para uma determinada situação
 - Por exemplo: quando alguma restrição de movimento da câmera ou da cena é conhecida a priori.

Conclusões

- Em relação às tendências de pesquisa na área, diversas direções podem ser enumeradas:
 - Ainda sobre o ponto de vista de estimação, pode-se citar como tópicos atuais de pesquisa o aprofundamento da abordagem direta, dada sua literatura reduzida, bem como acerca dos métodos híbridos**.

Conclusões

- Em relação às tendências de pesquisa na área, diversas direções podem ser enumeradas:
 - Já do ponto de vista sensorial, outros tipos de câmeras, sistemas multi-câmeras e fusão com outras modalidades sensoriais, devem constituir objeto de pesquisas futuras.
 - É importante ressaltar que um dos maiores desafios de pesquisa na área de SLAM se refere a reconstruções de grandes ambientes.

Divisão da apresentação

4. Sugestões de melhoria

Sugestões de melhoria

- A leitura do capítulo seria mais agradável se a definição de alguns termos próprios da área de Visão Computacional fosse definidos no capítulo, ou talvez no final do livro numa seção de “dicionário”.
- Da forma como se encontra pressupõe-se que o leitor já domina os termos utilizados nessa área. No caso do meu estudo tive que procurar uma quantidade considerável de definições fora do livro (Google).

Sugestões de melhoria

- Tomar o cuidado com a utilização de algumas palavras como “preciso”, “simples” que, no meu entendimento, não são adequadas no uso de determinados contextos do capítulo.
- O autor diz que a área é geralmente classificada de acordo com o tipo de filtragem apresentada, mas não explica o porquê decide abordar diferentemente a área em “primitivas geométricas” e “diretamente”.

Sugestões de melhoria

- Acredito que seria melhor, para ao capítulo, abordar também a questão da utilização de múltiplas câmeras no SLAM Visual e não apenas a monocular.
- A divisão do texto parece estar em forma de um trabalho acadêmico como Introdução, Fundamentação Teórica, Principais Abordagens e Conclusões. Não sei se é o adequado para o capítulo de um livro.

Sugestões de melhoria

- O autor diz: “Além disso, vários exemplos práticos, resultados experimentais, limitações e vantagens particulares de cada método são também discutidos.”
 - Eu não encontrei isso no capítulo, pelo contrário, acho que falta mais exemplos e resultados experimentais.

Referências Bibliográficas

- Esta apresentação é baseada no rascunho do capítulo “SLAM visual” do livro de Robótica do ICMC-USP.



FIM