



**Data Science
Academy**

www.datascienceacademy.com.br

Introdução à Inteligência Artificial

A Geometria dos Corpos

Nesse momento, assumimos que sabemos o que parece ser as partes do corpo da pessoa (por exemplo, sabemos a cor e a textura da roupa da pessoa). Podemos modelar a geometria do corpo como uma árvore de onze segmentos (braços e pernas esquerdos e direitos, respectivamente, tronco, rosto e cabelo no topo do rosto), cada um dos quais é retangular. Assumimos que, dada a pose do braço esquerdo, a posição e a orientação (pose) do braço inferior esquerdo é independente de todos os outros segmentos; que, dada a postura do torso, a posição do braço esquerdo é independente de todos os segmentos; e estender essas premissas de forma óbvia para incluir o braço direito, as pernas, o rosto e os cabelos. Tais modelos são chamados frequentemente de modelos de “pessoas de papelão”. Os modelos formam uma árvore, que normalmente é enraizada no torso. Vamos procurar a imagem que tenha a melhor correspondência para essa pessoa de papelão usando métodos de inferência para uma rede de Bayes de estruturada em árvore.

Existem dois critérios para avaliar uma configuração. Primeiro, um retângulo de imagem deve parecer com o seu segmento. Por enquanto, ficará vago o que isso significa exatamente, mas assumiremos que temos uma função ϕ_i que pontua o quão bem um retângulo de imagem corresponde a um segmento do corpo. Para cada par de segmentos relacionados, temos outra função que pontua o quão bem a relação entre um par de retângulos de uma imagem combina com as que são esperadas dos segmentos do corpo. As dependências entre os segmentos formam uma

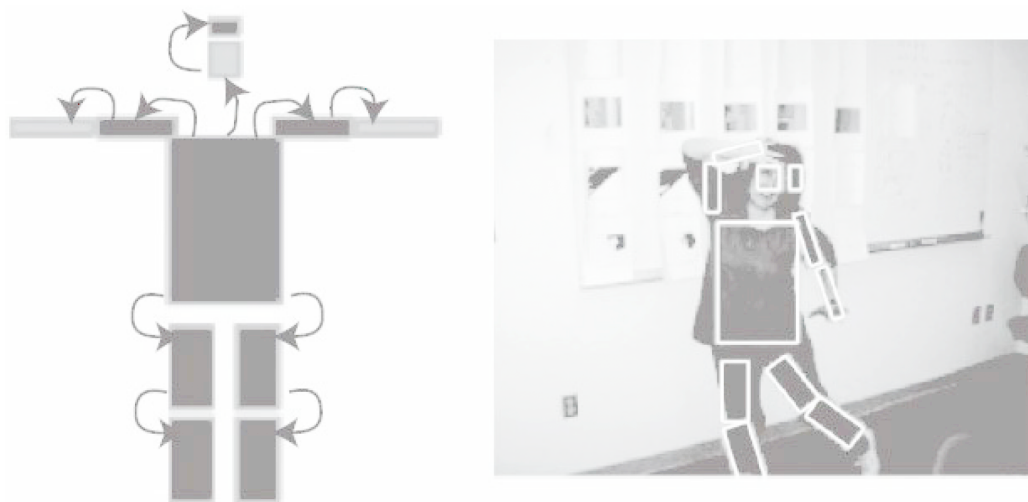
árvore, de modo que cada segmento tem apenas um pai, e poderíamos escrever $\psi_{i, \text{pa}(i)}$. Todas as funções serão maiores se a correspondência for melhor; assim, podemos considerá-las como sendo uma probabilidade de log. Então, o custo de uma combinação particular que aloca a imagem retângulo m_i para o segmento do corpo i é:

$$\sum_{i \in \text{segmentos}} \phi_i(m_i) + \sum_{i \in \text{segmentos}} \psi_{i, \text{pa}(i)}(m_i, m_{\text{pa}(i)})$$

A programação dinâmica pode encontrar a melhor correspondência porque o modelo relacional é uma árvore.

É inconveniente buscar em um espaço contínuo, e vamos discretizar o espaço da imagem dos retângulos. Discretizamos a localização e a orientação dos retângulos de tamanho fixo (os tamanhos podem ser diferentes para segmentos diferentes). Como os tornozelos e os joelhos são diferentes, precisamos distinguir entre um retângulo e o mesmo retângulo com um giro de 180°. Pode-se visualizar o resultado como um conjunto de pilhas muito grandes de pequenos retângulos de imagem, com corte em locais e orientações diferentes. Há uma pilha por segmento. Devemos agora encontrar a melhor alocação de retângulos aos segmentos. Isso é lento porque há muitas imagens de retângulos e, para o modelo que demos, escolher o tronco correto será $O(M^6)$, se houver M imagens de retângulos.

No entanto, várias acelerações estão disponíveis para uma escolha apropriada e o método é prático (figura abaixo). O modelo normalmente é conhecido como modelo de estrutura pictórico.



Um modelo de estrutura pictórica avalia uma combinação entre um conjunto de retângulos de imagem e uma pessoa de papelão (mostrado à esquerda), marcando a semelhança na aparência entre os segmentos do corpo e os segmentos de imagem e as relações espaciais entre os segmentos de imagem. Geralmente, uma combinação será melhor se os segmentos de imagem tiverem a aparência correta e estiverem no lugar certo com relação um ao outro. O modelo de aparência utiliza cores médias para cabelo, cabeça, tronco, braços e pernas. As relações relevantes são mostradas com setas. À direita, a melhor correspondência para determinada imagem obtida ao utilizar programação dinâmica. A combinação é uma estimativa razoável da configuração do corpo.



Detector de
caminhada lateral



Modelo de
aparência



Mapas de
parte do corpo

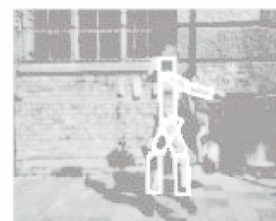
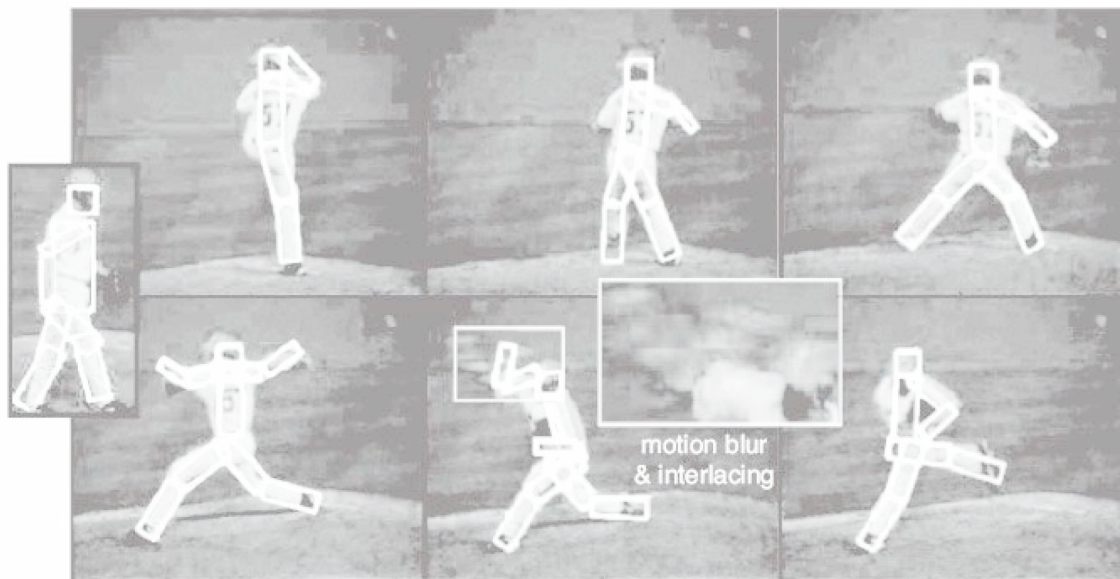


Figura detectada



Podemos rastrear o deslocamento de pessoas com um modelo de estrutura pictórica, primeiro obtendo um modelo de aparência, depois o aplicando. Para obter o modelo de aparência, digitalizamos a imagem para encontrar uma postura de caminhada lateral. O detector não precisa ser muito preciso, mas deve produzir poucos falsos positivos. A partir da resposta do detector, podemos ler os pixels que se encontram em cada segmento do corpo e outros que não se encontram sobre esse segmento. Isso torna possível construir um modelo discriminativo do aparecimento de cada parte do corpo, e eles são amarrados em um modelo de estrutura pictórica da pessoa que está sendo rastreada.

Finalmente, podemos traçar de maneira segura detectando esse modelo em cada quadro. À medida que os quadros na parte inferior da imagem sugerem, esse procedimento pode rastrear configurações de corpo complicadas, de mudanças rápidas, apesar da degradação do sinal de vídeo devido ao motion blur.

Lembre-se da nossa suposição de que sabemos o que precisamos saber sobre como a pessoa parece. Se estivermos combinando uma pessoa em uma única imagem, o recurso mais útil para pontuar combinações do segmento acaba por ser a cor. Características de textura não funcionam bem, na maioria dos casos, porque as dobras em roupas folgadas produzem padrões de sombreamento forte que sobrepõem a textura da imagem. Esses padrões são fortes o suficiente para interromper a textura verdadeira da roupa. Em trabalho recentes, reflete tipicamente a necessidade de as extremidades dos segmentos estarem razoavelmente próximas, mas normalmente não há restrições sobre os ângulos. Geralmente, não sabemos com o que uma pessoa se parece, e deve-se construir um modelo de aparências do segmento. Chamamos a descrição do que uma pessoa se parece como modelo de aparência. Se tivermos de relatar a configuração de uma pessoa em uma única imagem, poderemos iniciar com um modelo de aparência mal sintonizado, estimar a configuração com ele e então reestimar a aparência, e assim por diante. No vídeo, temos muitos quadros da mesma pessoa, e isso vai revelar a sua aparência.



Referências:

Livro: Inteligência Artificial

Autor: Peter Norvig