

Classificação Hierárquica de Caracteres Matemáticos



Breno Flesch Franco e Ricardo Alexandre Bastos

Orientadora: Profa. Dra. Nina S. T. Hirata

Instituto de Matemática e Estatística

Universidade de São Paulo

b flesch@linux.ime.usp.br, rabastos44@gmail.com



Introdução

O reconhecimento de caracteres em expressões matemáticas manuscritas é um grande desafio devido à inerente dificuldade de reconhecimento da escrita manuscrita e também à diversidade de possíveis caracteres.

O uso de uma abordagem hierárquica para o reconhecimento dos diversos caracteres presentes em uma expressão matemática [1] influenciou, neste trabalho, o desenvolvimento de um novo reconhecedor para o projeto ExpressMath. Além disso, modificações e adições foram feitas no software Math-Picasso, uma versão previamente desenvolvida no contexto do projeto ExpressMath, visando aperfeiçoar o que fora desenvolvido até então.

Reconhecimento de caracteres

Algoritmos de classificação que recebem como entrada um conjunto de características (quaisquer medidas que podem ser extraídas) de um símbolo manuscrito são utilizadas para o reconhecimento de caracteres. Os classificadores são geralmente “treinados” usando-se um conjunto de amostras classificadas manualmente.

Classificador Binário Hierárquico (BHC)

Para tratar adequadamente o grande número de classes, decidimos utilizar o Classificador Binário Hierárquico (BHC), que possui a capacidade de aprendizado em módulos e uma boa razão entre taxa de acerto e tempo despendido em treinamento e classificação.

O BHC [2] cria uma árvore binária onde:

- a raiz representa o conjunto de todas as classes, Ω ,
- cada nó interno representa uma metaclassa (um conjunto de classes) $\Omega_n \subset \Omega$, um extrator de características ψ_n e um classificador ϕ_n , possuindo dois nós filhos (que representam as metaclassas Ω_{2n} e Ω_{2n+1} , que compõem uma partição de Ω_n) e
- as folhas representam as classes $\omega_i (i = 1, 2, \dots, n)$.

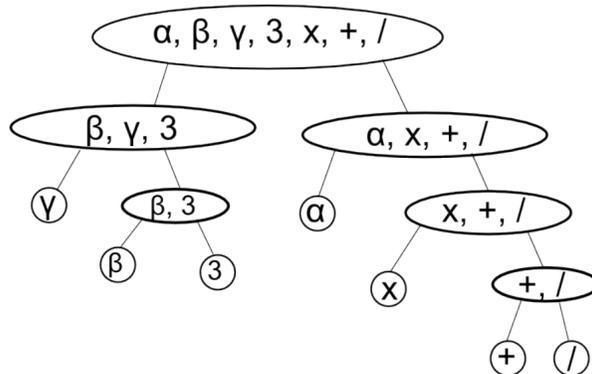


Figura 1: Exemplo de árvore de classificação gerada pelo BHC

Construção de um BHC

Um BHC pode ser construído usando técnicas de aprendizado computacional: para a divisão de um conjunto de classes Ω_n de um nó, utilizamos um algoritmo de aprendizado modular baseado em *clustering*, uma técnica que consiste no agrupamento de classes similares.

- Cada classe em Ω_n começa associada igualmente a duas metaclassas, Ω_{2n} e Ω_{2n+1} , com exceção da classe ω_1 , que é associada totalmente à metaclassa Ω_{2n} .
- A cada iteração, o grau de associação entre classes e metaclassas é atualizado utilizando-se um estimador de máxima verossimilhança da probabilidade de ω_i pertencer a Ω_{2n} , considerando-se a similaridade entre as características de ω_i e as das classes cuja associação a $\Omega_{2n} \geq 0.5$. Com isso, classes cujas amostras apresentam características mais próximas a ω_1 são associadas à mesma metaclassa.
- Calcula-se uma medida H_{ω_i} que expressa o quão divididas entre as duas metaclassas estão as amostras da classe ω_i .
- As associações e, em seguida, as medidas H são atualizadas até a medida H de todas as classes ficar abaixo de um limiar pré-definido, no qual se considera que cada classe está associada à metaclassa que apresenta maior verossimilhança.
- A partir daí, o conjunto pode ser dividido em dois e o algoritmo é aplicado recursivamente até que cada metaclassa contenha apenas uma classe.

O processo acima resulta em uma árvore binária tal que em cada nó interno está

associado o ponto médio e a matriz de covariância das amostras das classes que fazem parte da metaclassa correspondente ao nó. O ponto médio e a matriz de covariância associam uma distribuição de probabilidade ao nó.

Para classificar um símbolo, a árvore deve ser percorrida do nó raiz até um nó folha. O próximo nó a ser percorrido é decidido escolhendo-se o nó filho que apresenta maior probabilidade de conter o símbolo. A classe a que o símbolo pertence é aquela pertencente ao nó folha.

Resultados Parciais

As características escolhidas para a realização dos testes foram a contagem de laços e *quinas* (pontos onde o traço do caractere forma um ângulo agudo) do caractere, a distribuição dos pontos pelas seções de uma grade superposta ao caractere e três vetores que representam a direção média tomada em cada terço do caractere.

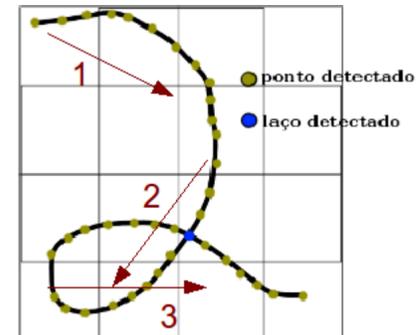


Figura 2: exemplo de características exibidas no caractere 2. Não há quinas, e um único laço.

Testes preliminares, com 15 classes de caractere e 20 amostras de treinamento por classe apresentaram taxas de acerto na classificação de um conjunto de testes com 4 amostras por classe entre 80% e 95%, o que sugere que o método é suficientemente robusto para o problema proposto.

Serão realizados testes mais abrangentes e mais rigorosos a fim de determinar o desempenho do classificador em relação a um maior número de classes.

Este trabalho é apoiado pelo CNPq.

Referências

- [1] C. P. Garcia. Uma abordagem hierárquica para o reconhecimento de caracteres em expressões matemáticas manuscritas. 2008.
- [2] S. Kumar. *Modular learning through output space decomposition*. PhD thesis, 2000. Supervisores: Ghosh, Joydeep e Crawford, Melba M.