

Proceeding Series of the Brazilian Society of Computational and Applied Mathematics

Redimensionamento de Descritores Baseados em Padrões N -ários Locais Generalizados para Classificação de Texturas

Pedro Matos da Silva ¹

Instituto Federal do Espírito Santo, Cariacica-ES

João Batista Florindo ²

Departamento de Matemática Aplicada - IMECC - Unicamp, Campinas-SP

1 Introdução

A classificação de imagens de texturas é sem dúvida um importante campo de pesquisa em visão computacional, com inúmeros trabalhos aplicados em diversas áreas da ciência. Grande parte dos métodos de classificação de imagens de texturas é baseada na extração de características locais, ou análise multiescala, ou ainda em uma combinação destas duas. O Padrão Local Binário (LBP) [1] é um dos métodos baseados em medidas locais mais bem sucedidos, juntamente com suas variantes. A ideia básica no código LBP de uma imagem em tons de cinza está na interpolação de P pontos em uma vizinhança circular de raio R em torno de cada *pixel*. Desta forma, para cada *pixel* x temos $LBP_{P,R}(x) = \sum_{i=0}^{P-1} s(g_i - g_x)2^i$, com g_x sendo a intensidade do *pixel* x , g_i a intensidade definida em cada ponto interpolado e $s(y) = 0$ se $y < 0$ ou $s(y) = 1$ se $y \geq 0$. É fácil ver que existem 2^P códigos possíveis para o $LBP_{P,R}$ assim definido, mas existem variações, como o $LBP_{P,R}^i$ e o $LBP_{P,R}^{iu2}$, cujo número de códigos é bem mais reduzido. A metodologia aqui proposta, além de se mostrar mais eficiente que as variações do LBP apresentadas, assim como outras da literatura, representa uma generalização do LBP, porém com parâmetro permitindo uma drástica redução da dimensão do descritor.

2 Metodologia e Resultados

Semelhantemente ao LBP, o que definimos aqui por Padrões Locais N -ários é obtido pela interpolação de $P = m \cdot p$ pontos em um círculo de raio r centrado em torno de cada *pixel* x . Para simplificar, este será representado por $LNP_{r,p,m}$. Com as coordenadas do *pixel* tais que $x = (x_1, x_2)$, os pontos interpolados serão dados por

$$y_i^k(x) = \left(x_1 + r \cos \left(\frac{2k\pi}{mp} + \frac{2i\pi}{p} \right), x_2 + r \sin \left(\frac{2k\pi}{mp} + \frac{2i\pi}{p} \right) \right),$$

¹pedroms@ifes.edu.br

²jbfiorindo@ime.unicamp.br

com $i \in \{0, 1, \dots, p-1\}$ e $k \in \{0, 1, \dots, m-1\}$. Além disso,

$$LNP_{r,p,m}(x) = \left(LNP_{r,p}^0(x), LNP_{r,p}^1(x), \dots, LNP_{r,p}^{m-1}(x) \right) \in \mathbb{Z}^m, \quad (1)$$

com $LNP_{r,p}^k(x) = \sum_{i=0}^{p-1} f\left(g(y_i^k(x)) - g(x)\right)N^i$, em que N é inteiro positivo, $g(x)$ é a intensidade do *pixel* x , $g(y_i^k(x))$ é um ajuste linear das intensidades dos 4 *pixels* mais próximos do ponto interpolado $y_i^k(x)$ e $\mathcal{IM}(f) = \{0, 1, \dots, N-1\}$.

Com esta definição, existem N^{pm} possibilidades para $LNP_{r,p,m}(x)$, que seria a dimensão máxima do descritor gerado. Embora cada *pixel* x gere pm pontos, ainda que todos ajudem na descrição, eles podem ser separados em m grupos, cada um com p pontos. Cada componente de $LNP_{r,p,m}(x)$ em ((1)) é formada por um grupo de p pontos, gerando um código para os padrões locais N -ários. Tomando-se como descritor o histograma dos valores de $LNP_{r,p}^k$, tal descritor terá dimensão igual a N^p .

Como exemplo, para tomarmos 16 pontos em torno de cada pixel, com $N = 4$, o histograma direto resultaria em um descritor com dimensão **4.294.967.296**. Fazendo-se $p = 4$ e $m = 4$ a dimensão seria reduzida para **256**. Observe-se que m não influencia na dimensão do descritor, mas sim na quantidade de vezes que cada *pixel* é “descrito”, ainda assim torna mais detalhada a descrição. Admitindo-se que a imagem tem tons que variam de 0 a M , o domínio de f será os inteiros de $-M$ a M . Particionando-se o intervalo $[-M, M]$ em N subintervalos I_0, \dots, I_{N-1} , pode-se definir $f(w) = j$, se $w \in I_j$.

Fazendo-se uma aplicação com $r = 1$, $N = p = m = 4$, a base UIUC teve **89.9%** de suas imagens classificadas corretamente e a base Outex, **82.35%**. Este é um resultado condizente e competitivo com o estado-da-arte na literatura. Nesta aplicação, os intervalos foram variáveis para cada imagem, $I_0 = [-255, -\mu]$, $I_1 = (-\mu, 0]$, $I_2 = (0, \mu]$ e $I_3 = (\mu, 255]$, em que μ é a variação modular média entre os *pixels* vizinhos. Para reduzir ainda mais o descritor, usa-se aqui estratégia semelhante ao $LBP_{P,R}^r$ de [1], reduzindo a dimensão do descritor para **70**.

Agradecimentos

Agradecemos o apoio financeiro da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) (Processo nº 2016/16060-0) e do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (Processo nº 301480/2016-8 e 423292/2018-8).

Referências

- [1] T. Ojala, M. Pietikäinen, e T. Mäenpää. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2002, (7) p.971-987.