

多尺度对称变换及其应用于定位人脸特征点

彭进业^{1,2}, 俞卞章¹, 王大凯², 李楠²

(1. 西北工业大学电子工程系, 陕西西安 710068; 2. 西北大学电子科学系, 陕西西安 710069)

摘要: 提出了一种在图像的反对称双正交小波分解数据域中, 实现多尺度对称变换的方法, 并将它应用于脸部图像中主要特征点的定位。这种方法与直接在原始图像数据域中的同类方法相比, 能大幅度减少计算量、加快定位速度, 并明显提高定位准确度。除此之外, 由于它是在小波分解数据域上实现的, 这就为在图像数据压缩域中提取人脸图像特征奠定了基础。

关键词: 对称变换; 特征点定位; 反对称双正交小波; 基于内容的图像检索

中图分类号: TP391 **文献标识码:** A **文章编号:** 0372-2112 (2002) 03-0363-04

Multi-Scale Symmetry Transform with Application to Location of Feature Points on Human Face Image

PENG Jin ye^{1,2}, YU Bian zhang¹, WANG Da kai², LI Nan²

(1. Department of Electronics, Northwestern Polytechnic University, Xi'an, Shaanxi 710068, China;

2. Department of Electronics, Northwest University, Xi'an, Shaanxi 710069, China)

Abstract: A method to implement multi scale symmetry transform in the wavelet coefficients domain decomposed by using anti symmetrical biorthogonal wavelets (ASBW) is presented with application to location of feature points on human face image. Compared to the original symmetry transform, which is performed on a single resolution level, our method can yield results with higher location accuracy without increment of computational burden. In addition, if an image is compressed based on ASBW decomposition, then our method provides a way to extract image features in compressed data domain without need of completely reconstructing the image. This is one of the most important technical consideration in the content based indexing (CBI) problem.

Key words: symmetry transform; location of feature points; anti symmetrical biorthogonal wavelets; content based indexing

1 引言

人脸识别是一种重要的个人身份鉴别方法, 近年来再次成为模式识别和人工智能领域的一个十分活跃的课题^[1]。而人脸图像中主要部件, 如眼、口等的定位是完成自动人脸识别任务的首要步骤。目前已提出了诸如马赛克方法^[2]、特征脸方法^[3]等将人脸视为一个整体模式的定位方法; 也提出了一些基于人脸个别器官的特征点的定位方法, 如: 神经网络方法^[4], 广义对称变换方法^[5]等。

Riesfeld 根据物体对称性提出的广义对称变换方法, 是实现对称物体定位的有效方法之一, 将其用于眼、口等人脸主要器官的定位, 取得了较好的实验结果, 并且具有对人脸轻微偏转、表情变化及光照变化等的鲁棒性^[5]。

记图像上任意一点 $p_k = (x_k, y_k)$, $k = 1, \dots, K$, 其中 (x_k, y_k) 为 p_k 点的坐标, K 为图像上像素总数。点 p_k 处的方向梯

度表示为

$$\nabla_{p_k} = \left[\frac{\partial}{\partial x} p_k, \frac{\partial}{\partial y} p_k \right]$$

并记

$$r_k = \log \left(1 + \|\nabla_{p_k}\| \right) \quad (1)$$

其梯度矢量相角为

$$\theta_k = \arctan \left(\frac{\frac{\partial}{\partial x} p_k}{\frac{\partial}{\partial y} p_k} \right) \quad (2)$$

对于图像上的任意两点 p_i, p_j , 其连线 l 与水平线的逆时针夹角记为 α_{ij} , α_{ij} 的取值为 $[0, \pi]$, 定义邻域

$$\Gamma(p) = \left\{ (i, j) \mid \frac{p_i + p_j}{2} = p \right\} \quad (3)$$

它是以 p 为连线中心点的点对集(见图1)。

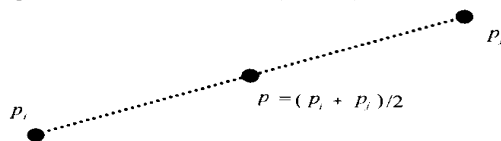


图1 p 为连线中心点的点对 p_i, p_j

定义 i, j 间高斯距离

$$D_o(i, j) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left(- \frac{\|p_i - p_j\|}{2\sigma} \right) \quad (4)$$

其中 σ 用于控制距离有效范围, 称为距离控制因子. 再记方向因子为

$$P(i, j) = [1 - \cos(\theta_i + \theta_j - 2\alpha_j)] [1 - \cos(\theta_i - \theta_j)] \quad (5)$$

则点 p 处广义对称变换的强度定义为

$$M_\sigma(p) = \sum_{(i,j) \in \Gamma(p)} C(i, j) \quad (6)$$

$$\text{其中 } C(i, j) = D_\sigma(i, j) P(i, j) r_{ij} \quad (7)$$

而广义对称变换的方向则定义为

$$\phi(p) = \max \left\{ \varphi(p) \mid \varphi(p) = \frac{\theta_i + \theta_j}{2}, (i, j) \in \Gamma(p) \right\} \quad (8)$$

通过计算图像各点的对称强度 $M_\sigma(p)$, 就可找出局部强对称点, 从而实现对称物体的对称中心准确定位, 并可根据 $\phi(p)$ 来确定其对称方向.

由式(7)、(8)可知, 对称变换的前提是计算图像每一象素点的方向梯度的模和相角. 但是, 由于图像中不可避免地含有一定的噪声干扰, 直接利用图像数据的差分来计算方向梯度是不可取的. 换言之, 在计算方向梯度之前对图像作某种平滑处理是必不可少的. 问题是采用怎样的平滑最为恰当, 因为过小的平滑窗口可能不足以消除噪声的不良影响, 而过大的平滑窗口将使方向梯度的局部变化趋于消失, 从而影响定位的精度. 因而, 应采用一种由粗至精的方法在分辨率上计算方向梯度和对称变换. 这就是多尺度对称变换的基本思想, 它与图像多尺度边缘提取的基本思想^[9]是一致的.

本文在提出多尺度对称变换的基本概念和实现方法(第2节)的基础上, 将它应用于人脸特征点定位并给出了实验结果(第3节), 最后第四节对全文作了简要总结与讨论.

2 基于反对称双正交小波的多尺度对称变换

小波变换的多尺度性质提供了在图像的多个分辨率上计算梯度的快速算法^[9], 但文献[6]中建议采用的小波函数属于二进小波, 不仅计算工作量较大, 而且其变换系数不利于直接用于编码压缩. 我们在[7]中证明了反对称双正交小波(ASBW)具有微分算子的功能. 事实上, 由于紧支撑且具有线性相位的反对称双正交小波的二尺度系数序列 $\{h_k\}$ 和 $\{g_k\}$ 的长度必均为偶数, 且两者相差为2的偶数倍^[8]. 因此, 不失一般性可令 $\{g_k\}$ 的对称中心为0.5, 长度为 $2K$, 起点下标为 $-K+1$, 终点下标为 K , 并且有

$$g_k = -g_{1-k}, \quad k = 0, 1, \dots, K \quad (9)$$

$$\text{考虑到 } \sum_{k=-K+1}^K g_k = 0 \quad (10)$$

二尺度关系可重写为

$$\begin{aligned} \phi(x) &= \sqrt{2} \sum_k g_k \phi(2x - k) = \sqrt{2} [g_{-K+1} \phi(2x + K - 1) \\ &\quad - \phi(2x + K - 2)] + (g_{-K+1} + g_{-K+2}) [\phi(2x + K - 2) \\ &\quad - \phi(2x + K - 3)] + \dots + (g_{-K} + g_{-K+1} + \dots + g_K) \\ &\quad \cdot [\phi(2x - K) - \phi(2x - K - 1)] \\ &= \sqrt{2} [g_{-K} \phi(2x + K - 1) - \phi(2x + K - 2)] \\ &\quad + (g_{-K+1} + g_{-K+2}) [\phi(2x + K - 2) - \phi(2x + K - 3)] \\ &\quad + \dots + g_{-K+1} [\phi(2x - K + 1) - \phi(2x - K)] \end{aligned} \quad (11)$$

令上式第 k 项的系数为 c_k

$$c_k = \sqrt{2} \sum_{i=-K+1}^k g_i, \quad k = -K+1, \dots, K-1 \quad (12)$$

则 c_k 是一个对称序列

$$c_k = c_{-k}, \quad k = 0, \dots, K \quad (13)$$

考虑到 $|g_k|$ 的值随 $|k|$ 的增大而减小及在式(12)的求和中正与负的抵销, 式(11)的各项中, 以 c_0 为系数的中间项将起主导作用. 也就是说, $\phi(x)$ 可近似表达为

$$\phi(x) \approx c_0 [\phi(2x) - \phi(2x - 1)] \quad (14)$$

如果 $\phi(x)$ 足够光滑, 根据中值定理, 式(14)可进一步表达为

$$\phi(x) \approx c_0 \frac{d}{dx} \phi(2x - 0.5) \quad (15)$$

这就是说, 任何紧支撑的反对称双正交小波 $\phi(x)$ 近似地正比于高一分辨率上与之相对应的尺度函数 $\phi(2x)$ 的导数. 但其精确程度可能各有不同, 这取决于式(15)的近似程度.

基于以上分析可以采用 Mallat 塔式算法对图像进行双正交小波分解并有

$$\left. \begin{aligned} d^{(1)}(x, y) &= f(x, y) * \phi_j(x) \phi_j(y) \propto \frac{\partial}{\partial x} (f(x, y) * \phi_j(2x) \phi_j(y)) \\ d^{(2)}(x, y) &= f(x, y) * \phi_j(x) \phi_j(y) \propto \frac{\partial}{\partial y} (f(x, y) * \phi_j(x) \phi_j(2y)) \\ d^{(3)}(x, y) &= f(x, y) * \phi_j(x) \phi_j(y) \end{aligned} \right\} \quad (16)$$

利用 $d^{(1)}(x, y)$ 和 $d^{(3)}(x, y)$ 的每一列作一次重构可得到

$$d_j^{(x)}(x, y) = f(x, y) * \phi_j(x) \phi_j(2y) \propto \frac{\partial}{\partial x} (f(x, y) * \phi_j(2x) \phi_j(2y)) \quad (17)$$

同理, 利用 $d^{(2)}$ 和 $d^{(3)}$ 的每一行作一次重构, 可得

$$d_j^{(y)} \propto \frac{\partial}{\partial y} (f(x, y) * \phi_j(2x) \phi_j(2y)) \quad (18)$$

我们将上述计算 $d_j^{(x)}$ 和 $d_j^{(y)}$ 的过程称之为“半重构”算法. 利用这一算法, 可以直接基于 ASBW 的分解数据计算多个分辨率上近似图像的方向梯度的模值

$$\|\nabla_j(x, y)\|^2 = (d_j^{(x)})^2 + (d_j^{(y)})^2$$

和相角 $\theta_j(x, y) = \arctan(d_j^{(y)}(x, y) / d_j^{(x)}(x, y))$

从而为实现多尺度对称变换奠定基础.

值得注意的是当进行多尺度对称变换时, 距离控制因子 σ 也应随着分辨率级别的变化而变化, 即

$$\sigma_j = \sigma_0 2^{-j} \quad (19)$$

在小波分解数据域中实现对称变换与在原始图像数据上的广义对称变换相比, 其突出的优点是: 可充分利用小波变换的多分辨率特性, 即先在低分辨率上实现快速粗定位, 再根据粗定位的结果在高一级分辨率所对应的一个邻域内进行强对称点搜索, 从而逐级提高定位精度. 这样不仅大大减少全过程的计算工作量, 并且由于低分辨率图像只保存了较强的大结构对象的梯度信息, 因而本方法对噪声有很强的抑制作用. 除此之外, 基于 ASBW 的分解数据还可以直接用于图像编码压缩, 所以这一方法使得在图像压缩数据域中提取基于方向梯度的图像特征成为可能, 对于在压缩数据域中实现基于内容的图像检索具有重要意义.

3 人脸正面图像中眼与口的定位方法

作为多尺度对称变换的一个具体应用,本节讨论利用它实现人脸图像中许多重要特征点(如眼、口等)的定位问题。

3.1 眼部定位

在采用小波域对称变换进行人脸定位时,首先应根据待定位特征点对称区域特性选择恰当的距离控制因子 σ ,用于人眼定位时,就要考虑眼球和眼白部分的大小和形状。

通过对眼睛简化模型(如图2)的分析,我们将眼部划分为圆形对称区(瞳孔区)和椭圆形对称区(眼帘之内),因此我们将高斯距离因子也分为两段,表示为

$$D_d(i, j) = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma_0} \exp\left[-\frac{(x_j - x_i)^2 + (y_j - y_i)^2}{2\sigma_0^2}\right], & r < \sigma_0 \\ \frac{1}{\sqrt{2\pi}\sigma_1\sigma_2} \exp\left[-\frac{(x_j - x_i)^2}{2\sigma_1^2}\right] \exp\left[-\frac{(y_j - y_i)^2}{2\sigma_2^2}\right], & \text{其它} \end{cases} \quad (20)$$

其中 $r = (x_j - x_i)^2 + (y_j - y_i)^2$,根据我们目前应用场合(证件人面图像),我们取 σ_0 等于各级近似图像宽度的1/20,以及 $\sigma_1 = 3\sigma_0$, $\sigma_2 = 1.5\sigma_0$ 。另外,在圆形区域内部作对称变换时,将方向权重因子 $P(i, j)$ 修改为

$$P(i, j) = [1 + \cos(\theta_i - \alpha_j) \cos(\theta_j - \alpha_j - \pi)] \left| \alpha_j - \frac{\pi}{2} \right| \quad (21)$$

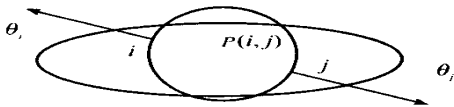


图2 眼睛简化模型

其理由是这样的,对于圆形边缘上的两对称点 i, j ,其梯度方向 θ_i, θ_j 与 i, j 连线成 0° 或 180° ,因此,因子 $[1 + \cos(\theta_i - \alpha_j) \cos(\theta_j - \alpha_j - \pi)]$ 的设置能增强对圆形物体如眼球的检测,而 $|\alpha_j - \pi/2|$ 的引入是考虑到眼球可能部分地被上、下眼皮所覆盖,越是水平方向对称点其权重应越大,以减弱垂直方向的强对称性所产生的强对称点。

3.2 口部定位

根据眼睛的定位结果和眼部与口部的几何关系,能粗定口部区域,这就有效地减少了对称变换的范围。但由于口部的变化比眼部的变化大,因而要实现口部的精确定位比较困难。为了提高口部定位精度,我们根据一般口型特点,将高斯距离因子修正为椭圆形区域,表示为

$$D_o(i, j) = \frac{1}{\sqrt{2\pi}\sigma_1\sigma_2} \exp\left[-\frac{(x_j - x_i)^2}{2\sigma_1^2}\right] \exp\left[-\frac{(y_j - y_i)^2}{2\sigma_2^2}\right] \quad (22)$$

取 σ_2 等于各级近似图像宽度的1/10,并取

$$\sigma_1 = c \sigma_2$$

式中 c 为常数。根据口的张合的程度来设置,一般地随着口的张合 c 的取值范围为1至6。

另外,由于嘴唇形状的变化、牙齿的显露等情况将影响口部的上下对称性,但对左右对称性的改变不大,因此我们对方向权重因子修正为

$$P(i, j) = [1 - \cos(\theta_i + \theta_j - 2\alpha_j)] [1 - \cos(\theta_i - \theta_j)] \left| \alpha_j - \frac{\pi}{2} \right| \quad (23)$$

式中 $|\alpha_j - \pi/2|$ 因子的引入增加了水平方向对称点的权重。

3.3 实验

利用上述方法我们对人脸图像眼部和口部的定位进行了实验研究,并与广义对称变换结果进行了比较。实验中选用的反对称双正交小波是 MATLAB 中的 bior1.5。

我们首先选用 MIT 人脸库中的 16 人(如图3),每人 3 幅,共 48 幅图像(其尺寸为 240×256)作为实验数据库,因为这些图像中包含表情、光照的一定变化,还有些戴眼镜,有些有浓密的胡须,有些还有稍微偏转。



图3 MIT 图库中的 16 人

先对整幅图像作 2 级小波分解,再在各级细节系数上作对称变换来定位眼睛中心点,实验中选择待选强对称点数目为 4。图4是其中一幅图像在 $j = -1$ 和 $j = -2$ 级分辨率下的眼部定位例图,用十字型标记。同时,我们也采用广义对称变换方法^[5]对这 48 幅图像作了在原图上的眼部定位实验。表1中列出了分别采用本文提出的多尺度对称变换和文^[5]中原图上广义对称变换进行人脸眼部定位的实验时间比较,以及各级分辨率上及原图上广义对称变换的定位准确度(定位在眼球半径的1/2内就认为是准确的)比较。

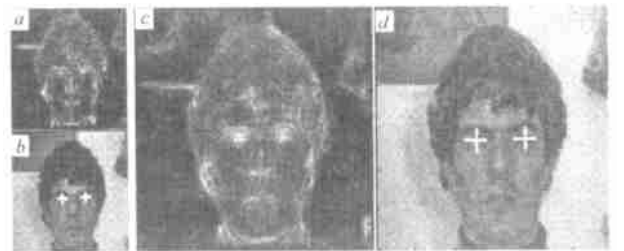


图4 多尺度对称变换用于眼睛定位 (a) $j = -2$ 对称强度 (b) $j = -2$ 定位结果, (c) $j = -1$ 对称强度, (d) $j = -1$ 定位结果

我们在实验中还发现,多尺度对称变换对戴眼镜情况下的定位能取得好的定位结果。在实验图像中,除镜片反光特别严重的一幅以外,其余都取得了较为满意的定位结果,如图5所示。在这一点上,它也明显优于在原图上的广义对称变换。分析其原因,在于小波变换降低了分辨率,减弱了镜片反光对定位的影响。

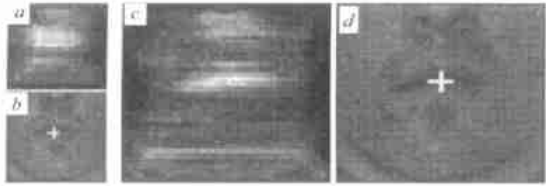


图5 多尺度对称变换($j = -1$)用于戴眼镜时定位结果

表1 定位方法的实验时间(采用相对时间)及准确度比较

	小波域对称变换		原图上广义 对称变换
	$j = -2$	$j = -1$	
定位时间	0.006	0.065	1
准确度	87.5%	93.8%	81.3%

根据眼睛的定位结果,由眼部与口部的几何关系,可有效地减少口部定位的搜索范围.除有浓密胡须的一幅定位不够准确以外,其它都取得了好的定位结果.在 $j = -1$ 级分辨率上的定位准确度达 93.8%.图 6 是其中一幅图像在 $j = -1, j = -2$ 级下的定位结果(取 $c = 6$).

图6 多尺度对称变换用于口部定位 (a) $j = -2$ 对称强度, (b) $j = -2$ 定位结果, (c) $j = -1$ 对称强度, (d) $j = -1$ 定位结果

为了验证本文方法对面部比较平坦的人种的有效性,我们在实验室环境下拍摄了多幅汉族人脸照片($256 \times 256 \times 8\text{bit}$ 的灰度图像),面部表情和光照条件有较明显的变化.实验方法与上相同,眼部、口部强对称待选点数目均为 4.图 7 是其中一人的图像在 6 种情况下的定位结果($j = -1$ 级).

图7 自拍图像在 $j = -1$ 级下的定位结果

从图 7 可见,在光照等外部环境变化及表情等内部因素变化的多种情况下均取得了较满意的定位结果.其中眼部定位只有 c 图由于眼睛完全闭合而使强对称点稍微靠近眉毛方向.而口部定位在口型与光照有较大变化的情况下,仍然取得了十分理想的定位结果.

4 结论与讨论

基于反对称双正交小波的微分功能,本文提出了一种在小波分解数据域上实现对称变换的方法.这种变换可以在不同分辨率下求取对称性.以此为基础,提出了一种新的人脸定位方法,可以在某一分辨率下独立求取对称性实现定位;也可以先在较粗一级定位,再在高一级分辨率下局域地作对称变换,来进一步提高定位速度和精度.通过对高斯距离因子、方

向因子的修正和距离控制因子的估计,求取包括人眼中心点在内的强对称点,最后通过对待选点的筛选来得到人眼的中心点.根据眼睛定位结果,可以初步确定口部区域.而针对口部特点选用求椭圆域对称中心方法,因而实现了口部的高准确度快速定位.

与文献[5]中方法比较,本方法不仅提高了定位准确度,大幅度减少计算量.而且由于它能在小波分解数据域实现图像定位,从而为进一步在图像压缩数据域中提取图像特征奠定了基础.

参考文献:

- [1] Chellappa R, Wilson C L, Siroshy S. Human and machine recognition of faces: a survey [J]. Proceedings of the IEEE, 1995, 83(5): 705–740.
- [2] Yang G Z, Huang T S. Human face detection in a complex background. Pattern Recognition [J]. 1994, 27(1): 53–63.
- [3] Turk M, Pentland A. Eigenfaces for recognition [J]. Cognitive Neuroscience, 1991, 3(1): 71–79.
- [4] Lin S H. Face recognition/detection by probabilistic decision based neural network [J]. IEEE Trans. Neural Networks, 1997, 8: 114–132.
- [5] D Reisfeld, H Wolfson, Y Yeshurun. Context-free attentional operators: the generalizes symmetry transform [J]. Computer Vision, 1995, 14: 119–130.
- [6] Mallat S, Zhong S. Characterization of signals from multi scale edges [J]. IEEE Trans. 1992, PAMI 14(7): 710–732.
- [7] HongBing Liu, Jinye Peng, Dakai Wang. Multiscale edge detection by using anti-symmetrical biorthogonal wavelets [A]. 3rd International Conference on Information, Communications and Signal Processing [C], Singapore, 2001: P0296. pdf.
- [8] M Vetterli, C Herley. Wavelets and filter banks: theory and design [J]. IEEE Trans., 1992, SP 40(9): 2207–2232.

作者简介:



彭进业 男. 1964 年 5 月出生于湖南涟源. 1996 年毕业于西北大学电子系, 获理学硕士学位, 毕业后留校任教至今. 现为西北大学电信工程教研室主任, 副教授, 西北工业大学博士研究生. 在国内外学术刊物发表论文近 20 篇. 主要研究兴趣包括数字图像处理、模式识别、小波应用研究.



俞卡章 男. 教授, 博士生导师. 主要研究领域为: 信号处理、图像处理、模式识别.