

基于复数小波能量特征和支持向量机的 图像匹配算法

王俊卿 黄莎白 史泽林

(中国科学院沈阳自动化研究所, 沈阳 110016)

摘要 为了对图像中发生平移、伸缩及旋转等形变的目标进行有效检测和跟踪,提出了基于复数小波能量特征和支持向量机的图像匹配算法,以便把图像匹配问题转化为图像分类问题。该算法首先利用复数小波的方向选择性、多尺度特性及近似平移不变性来抽取图像能量的均值、均方差及偏度等统计特征,并将其作为支持向量机的输入参数,用于训练模板样本集合,以获得支持向量,然后对由输入图像构成的与模板大小相同的所有子图像进行测试。这是一个粗精结合的两步算法,即先运用支持向量机筛选出候选目标集合,再运用非线性距离判优准则来确定检测出的候选目标图像集合中的最优匹配。实验结果表明,该算法克服了传统图像匹配方法搜索目标时存在的置信度问题,通过与基于径向基的神经网络学习方法和基于灰度相关的匹配方法比较可见,该算法在性能上优于这两个方法,并能得到满意的匹配结果。

关键词 图像匹配 复数小波 支持向量机 平移不变性 方向选择性

中图分类号: TP391.41 文献标识码: A 文章编号: 1006-8961(2004)09-1075-05

An Algorithm of Image Matching Based on Both Complex Wavelet Energy and SVM

WANG Jun-qing, HUANG Sha-bai, SHI Ze-lin

(Shenyang Institute of Automation Chinese Academy of Sciences, Shenyang 110016)

Abstract This algorithm, which is based on both statistical characteristics of complex wavelet energy and SVM, is proposed in order to effectively detect and track targets in image, which may cause changes, such as translation, scaling and rotation. So, the problem of image matching is transformed as that of classification. The transformation of complex wavelet that has properties of scale, shift invariant and directional selectivity effectively extract the statistical characteristics of image, such as mean, standard deviation and skew. The statistical characteristics of sample templates are input into SVM to train support vectors of SVM. Then, those statistical characteristics of any sub-image from original image are input into SVM in order to match target. This is a two-stage algorithm of coarse-to-fine. Firstly, the set of candidates is sifted by SVM. Secondly, a new optimal rule, which is nonlinear distance function, is proposed to decide the optimal matching from the candidate set. Those experimental results show that this algorithm addresses the problem of confidence level, which generally exists in traditional matching methods. This algorithm's performance is superior to those of both learning method of neural network based on RBF and gray-level correlation matching method, which compares with them. Finally, a good matching result is obtained.

Keywords image matching, complex wavelet, support vector machine (SVM), shift invariant, directional selectivity

1 引言

众所周知,图像匹配在人脸识别、指纹识别、图

像检索、交通监控、成像制导等领域有着广泛的应用,而传统的图像匹配方法主要是先通过提取图像的形状特征、纹理特征或灰度特征,并将其作为图像匹配的特征向量,然后采用距离测度来比较图像间

的相似性,以决定图像是否匹配。在特征空间中,对于分类问题,以择近原则作为相似测度不是所有情况都适用,因为单纯通过距离测度不能区别两类不同物体。对于图像匹配,也存在类似问题。

本文提出的图像匹配算法是把模式匹配问题看作模板图像类和非模板图像类的两分类问题,即通过学习样本,匹配相似目标子图像。其匹配是首先采用复数小波抽取图像的纹理特征;然后采用支持向量机(support vector machine, SVM)对给定的两类样本的纹理特征进行学习;接着通过把搜索的目标图像所包含的所有子图像作为测试图像来得到与模板图像相似的候选图像集;最后根据本文提出的最优判定准则进行后处理,即可匹配到最相似的目标子图像。

2 复数小波、支持向量机的基本理论与最优准则确定

2.1 复数小波

众所周知,有效的纹理特征表示对于图像匹配是重要的,而局域化分析的纹理特征则主要是通过 Gabor 变换^[1,2]和实数离散小波变换(discrete wavelet transform, DWT)^[3,4]得到。其中 Gabor 滤波器与人的视觉系统有许多相似的特性^[5],并可提供方向选择特性和多尺度的空间频率特性,其主要不足是计算量大;DWT 在数据压缩方面虽有很好的应用,但用其提取图像纹理特征来进行图像匹配或检索的效果并不特别理想,其主要原因是 DWT 的小波系数对平移变化敏感,不仅在每个尺度下只有 3 个方向子频带(水平、垂直、对角线方向),并且在对角线方向不能区分 $\pm 45^\circ$ 方向的频带。

复数离散小波变换(complet DWT, CDWT)相对于传统实数离散小波变换,其优点是它有近似的平移不变性,且在不同尺度下有 6 个可分的正负方向频带($\pm 15^\circ, \pm 45^\circ, \pm 75^\circ$)^[6,7],其中 6 个滤波器在频域的有效轮廓范围示意图如图 1 所示。

其相对于 Gabor 变换主要是速度优势,同时具有类似 Gabor 变换的特性,且比 Gabor 变换有更多的冗余。二维 CDWT 采用 Mallat 算法,可通过一组复数滤波器来分别在行列方向进行滤波。

算法中采用的复数小波变换的低通与高通滤波器系数分别为:

$$h_0 = [1 - \mathbf{j}, 4 - \mathbf{j}, 4 + \mathbf{j}, 1 + \mathbf{j}]/10$$

$$h_1 = [-3 - 8\mathbf{j}, 15 + 8\mathbf{j}, -15 + 8\mathbf{j}, 3 - 8\mathbf{j}]/48$$

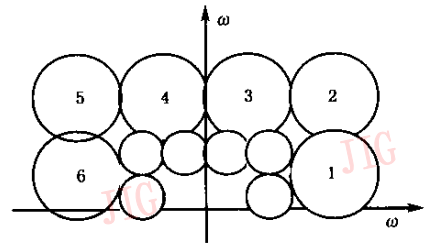


图 1 每个尺度下不同子频带方向分布示意图

对图像进行 CDWT 变换,每个尺度都可分别分解出由实部系数构成的 6 个不同方向的子图像和由虚部系数构成的 6 个不同方向的子图像。通过对这些小波系数进行归一化,即可消除图像对比度及目标变化尺度对匹配的影响。根据同一滤波器生成的与复数实虚部对应的子图像,即可求各方向子频带的能量子图像,其产生的 6 个能量子图像可表示成 $D_{m,k}$,其中 $k=1, \dots, 6$ 对应 6 个方向子频带, $m=1, 2, \dots$ 对应 CDWT 变换的级数。定义 $E(m,k), \mu(m,k), \sigma(m,k), S(m,k)$ 分别对应 $D_{m,k}$ 的能量(式(1))、能量均值(式(2))、能量均方差(式(3))、能量偏度(式(4))。

$$E(m,k) = \sum_{j=1}^N \sum_{i=1}^M |D_{m,k}(i,j)| \quad (1)$$

$$\mu(m,k) = \frac{\sum_{j=1}^N \sum_{i=1}^M |D_{m,k}(i,j)|}{M \times N} \quad (2)$$

$$\sigma(m,k) = \sqrt{\frac{\sum_{j=1}^N \sum_{i=1}^M (|D_{m,k}(i,j)| - \mu(m,k))^2}{M \times N - 1}} \quad (3)$$

$$S(m,k) = \frac{1}{M \times N} \sum_{j=1}^N \sum_{i=1}^M \left(\frac{|D_{m,k}| - \mu(m,k)}{\sigma} \right)^3 \quad (4)$$

其中, M, N 分为图像的行数和列数。若把能量均值、能量均方差和能量偏度作为匹配的特征向量,再用算法对图像进行了两级 CDWT 分解,则由式(1)、式(2)、式(3)及式(4)可得 36 个特征,即 $(\mu(1,1), \sigma(1,1), S(1,1), \mu(1,2), \sigma(1,2), \dots, \mu(1,6), \sigma(1,6), S(1,6), \mu(2,1), \sigma(2,1), S(2,1), \mu(2,2), \sigma(2,2), \dots, \mu(2,6), \sigma(2,6), S(2,6))$,而这些特征就可作为支持向量机的输入参数。

2.2 支持向量机基本理论

支持向量机基本思想是由线性可分情况下的最优分类面发展而来的^[8,9](如图 2 所示)。

图 2 表示两类样本, H 为分类面, H_1, H_2 分别

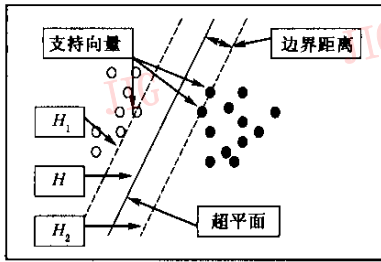


图 2 最优超平面

为过各类中的离 H 最近的样本,且平行于分类面 H 的平面,它们之间的距离叫做分类间隔 (margin)。设线性可分样本集为 $(x_i, y_i), i = 1, \dots, n, x_i \in R^d, y_i \in \{+1, -1\}$ 是类别标号。设 $x \in R^d, w$ 为权向量, b 为分类阈值,则其线性判别函数的一般形式为 $g(x) = w \cdot x + b$, 分类面 H 的方程为 $w \cdot x + b = 0$, 对它进行归一化,以使得两类所有样本都满足 $|g(x)| \geq 1$, 即,使离分类面最近样本的 $|g(x)| = 1$, 这样分类间隔就等于 $2 / \|w\|$, 可见,使间隔最大,等价于使 $\|w\|$ 最小;而要求分类面对所有样本正确分类,就是要求它满足

$$y_i [(w \cdot x_i) + b] \geq 1, i = 1, \dots, n \quad (5)$$

在满足约束条件(式(5))情况下,使下式

$$\Phi(w) = \frac{1}{2} \|w\|^2 \quad (6)$$

成立的最小的分类面叫做最优分类面, H_1, H_2 上的训练样本点就称做支持向量。由于将最优分类面转化为求 $\|w\|$ 的最小值问题,可通过 Lagrange function 求极值方法解决,从而可求出如下最优分类超平面 (optimization separation hypersurface, OSH)。

$$\begin{aligned} \hat{w} &= \sum_{i=1}^l a_i y_i x_i \\ \hat{b} &= -\frac{1}{2} (x_r + x_s) \end{aligned} \quad (7)$$

其中 $a_i > 0$ 为 Lagrange 系数, x_r, x_s 分别是标记 +1 的对应类和标记 -1 的对应类的第 r 个和第 s 个支持向量,而 $y_r = +1$ 为 x_r 的类标记, $y_s = -1$ 为 x_s 的类标记, \hat{w} 是 OSH 的权向量, \hat{b} 是其分类阈值。最优分类函数为

$$f(x) = \text{sign} \left(\sum_{i=1}^n a_i y_i (x_i \cdot x) + \hat{b} \right) \quad (8)$$

将 OSH 推广到非线性可分情况,可引入松弛变量 ξ_i 和惩罚因子 c 。广义 OSH 最小化式为

$$\Phi(w) = \frac{1}{2} \|w\|^2 + c \sum_{i=1}^n \xi_i \quad (9)$$

其中 c 是常量。这个最优问题可转化为用 Lagrange 求极值的方法解决,但需增加对 Lagrange 乘子的约束即 $0 \leq a_i \leq c, i = 1, \dots, n$ 。尽管通常遇到的分类问题是线性不可分的,即样本在输入空间不能被任何线性超平面分开,但在非线性映射的高维特征空间却可能是线性可分的,因为 SVM 可以在这个高维空间中构成一个 OSH。根据 Mercer 定理,无需直接计算这个映射,而是通过用核函数 $K(x_i, x_j)$ 代替 $x_i \cdot x_j$ 来完成非线性映射,这样此方法就解决了维数灾难问题^[10]。对于给定的核函数,则最优分类函数(式(8))可修改为

$$f(x) = \text{sign} \left(\sum_{i=1}^n a_i y_i K(x_i, x) + \hat{b} \right) \quad (10)$$

2.3 最优准则函数确定

将目标图像中与模板大小相同的所有子图像作为 SVM 的测试集,经过支持向量机分类后,产生的与模板图像相似的候选子图像集合用 S 表示。因为最优分类面存在线性可分和线性不可分两种情况,所以经过支持向量机分类得到的结果 S 是一个粗匹配结果,如果 S 非空,则需要对其进行后处理。

在线性可分情况下,可通过优化式(6)来得到最优分类面,以便把两类样本正确分开。由式(5)可见,由于 S 中子图像到最优超平面距离都大于等于 1,此时 S 中的子图像都是正确匹配的目标,所以后处理只是在这 S 中寻找最优匹配。

在线性不可分情况下,则需要通过优化式(9)来得到最优分类面。因为引入了松弛变量和惩罚因子,所以存在错分的可能,即存在候选子图像到超平面距离小于 1 的可能,所以要分情况考虑 S ,如需首先考虑 S 中到超平面的距离都大于等于 1 的子图像,由于这些子图像都是正确匹配的目标,因此可从这些子图像集合中寻找最优匹配,如果 S 中没有距离大于等于 1 子图像,则需找到超平面距离最大的子图像;如果其距离大于预先设定的阈值,则将其看作最优匹配的结果,否则认为当前搜索图像中无匹配目标。

在粗匹配过程中,不但极大缩小了候选子图像的范围,而且解决了传统图像匹配方法存在的置信度问题。此时,如果检测结果 S 为空,则说明在这个搜索的目标图像中无匹配目标;如果 S 非空,根据上述分析,则需要利用一个判定准则来确定 S 中与模板图像最相似的目标子图像。本文提出了一个新的判优准则函数,这是算法的精匹配过程,该策略不仅增强了算法的鲁棒性,而且提高了匹配的精度。

在二维情况下,点 (x_0, y_0) 到直线 $Ax+By+C=0$ 距离是 $\frac{|Ax_0+By_0+C|}{\sqrt{A^2+B^2}}$,同理,根据式(10)候选子图像的特征向量 x_i 到超平面的距离可以定义如下:

$$d(x_i) = \frac{\left| \sum_{j=1}^n a_j y_j K(x_j, x_i) + \hat{b} \right|}{\left\| \sum_{j=1}^n a_j y_j x_j \right\|} \quad (11)$$

式中, x_j 为模板类的支持向量, n 为该类支持向量的数目。一般式(11)可作为最优准则函数,但在 S 非空条件下,应分以下两种情况考虑最优准则判定函数:如果 S 中存在 x_i ,其 $d(x_i)$ 大于等于1,则采用K-近邻法计算 S 的聚类中心,并找 S 中离聚类中心最近的特征向量 x_k ;否则计算

$$x_k = \operatorname{argmax} \{d(x_i) > T, i = 1, \dots, L, \text{ 且 } 0 < d(x_i) < 1\} \quad (12)$$

其中, x_i 为候选图像特征向量, L 为 S 大小, T 为预先设定的阈值, x_k 为最优匹配目标子图像的特征向量。

3 算法详细描述

(1) 首先模板图像进行旋转、缩放、平移及对比度的变换,其中这些变换的比例系数是通过随机数发生器产生,该随机数是在预先设定的变化范围内,该过程相当于对模板图像进行若干次仿射变换,其模板变换后的样本集的分类标志 $y=+1$;然后把包含其他各种景物的图像作为非模板图像样本集,其分类标志为 $y=-1$;最后把这两类样本集作为SVM的训练样本。

(2) 对训练样本图像进行复数小波变换,并通过式(2)、式(3)、式(4)提取能量特征(均值、均方差、偏度)作为SVM的输入参数,再采用SMO(sequential minimal optimization)算法^[11]产生支持向量。

(3) 把搜索的目标图像包含的与模板图像大小相同的所有子图像作为SVM的测试集合来进行复数小波变换,同时用式(2)、式(3)、式(4)提取能量特征(均值、均方差、偏度),并根据式(10)进行分类,以产生与模板图像相似的候选图像集合 S ,如果 $S=\emptyset$ (\emptyset 表示空集),则说明该图像中无匹配目标,算法结束,否则转到步骤(4)。

(4) 对 S 中的子图像利用2.3节提出的最优判定准则函数进行最优判决来找到最优匹配的目标子图像。

4 实验结果、参数估计方法及性能分析

4.1 实验结果

实验中选用的训练样本集大小是2100,训练采用高斯RBF(radius base function)核函数,样本图像是 48×48 大小的256灰度级图像。模板图像是包含两个小人的 48×48 大小的灰度图像,模板图像集合是根据仿射变换,采用随机数发生器随机产生变换矩阵和偏移量获得的,共产生了1050个模板子图像。SVM算法采用Platt提出的SMO算法来训练支持向量^[11],经过实验,取 $\sigma^2=90, c=3$,即可使该算法得到理想的实验结果。在该实验中得到463个支持向量, $b=0.909676$ 。利用该算法对50帧序列图像进行试验,正确匹配率达到90%。图3列出其中6幅匹配图像。

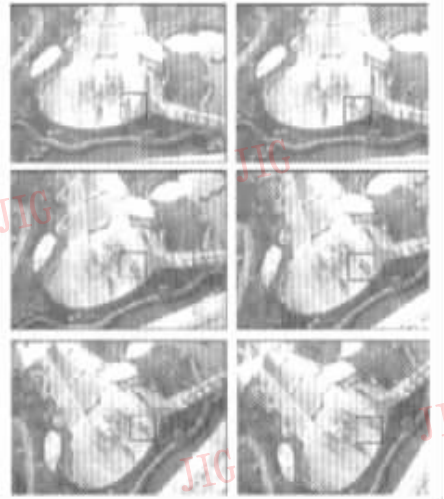


图3 6幅匹配结果图像

4.2 参数估计方法

在非线性可分情况下,参数 c 决定了分类边界错误率,这种错误既包括误分的样本,又包括位于类边界与最优分类面之间的正确样本。若 c 越小,则两分类面之间的距离越大,反之越小。当分类面之间的距离归一化后, c 的大小变化对SVM的影响表现为错分样本数目的多少,即 c 越小,错分样本可能越多,反之越少。根据该性质,可结合样本学习速度来取适当 c 值,而对 σ^2 进行估计,则需采用以下两个约束条件作为评价准则:一是根据分类器错误率最小,二是根据训练样本到最优分类面的平均距离与1的接近程度,通常可采用C-法对分类器性能进行评估。综合以上评价准则,通过试探法即可找到近似

最优的 c 和 σ^2 。实际上,对参数 c 和 σ^2 的确定是一个非常值得研究的问题。

4.3 性能分析

为说明算法的性能,还分别与方法 1(采用径向基神经网络作为分类器,输入特征量与 SVM 相同)和方法 2(传统的基于单模板的灰度相关法)进行了比较,3 种方法都在同一平台下运行。表 1 列出 3 种方法主要性能的比较。

表 1 3 种方法的性能比较

方法	每帧匹配时间(s)	C-法正确率(%)	匹配正确率(%)
本文算法	2.6	100	90
方法 1	3.1	100	78
方法 2	0.41		40

由于分类器是通过离线学习完成,所以样本测试时间可作为匹配时间。从不同算法比较结果看,本文算法性能比基于径向基神经网络(方法 1)好;相关算法(方法 2)运行效率虽高于本算法,但匹配正确率却不高,同时还存在置信度问题;本算法匹配时间由特征抽取时间、测试时间、最优匹配时间 3 部分构成,其中特征抽取处理耗时最多,但由于复数小波变换易于并行实现,所以计算量大的问题可通过并行方法和硬件电路或高速处理器件解决。

5 结论

本文提出的算法特点是结构清晰(主要分为复数小波特征提取阶段、SVM 学习分类阶段及最优目标定位阶段)、运行速度相对径向基神经网络快及匹配正确率高。该算法不仅解决了在传统图像匹配方法中搜索目标存在的置信度问题,而且对目标在背景中平移,旋转,伸缩等变化的适应能力强,且特征提取阶段易于并行实现。但该算法也有不足需要改进,如 SVM 核函数参数估计及计算量相对较大等问题,因此需要对参数估计进行研究和对算法进行优化,以提高匹配效率,另外对训练样本的设计也是提高正确匹配率的有效手段。总之,该算法在机器学习领域对研究图像匹配是一个非常可取的方法。

参考文献

- Manjunath B S, Ma W Y. Texture features for browsing and retrieval of image data [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1996, 18(8): 837~841.
- Ma W Y, Manjunath B S. Texture features and learning similarity [A]. In: Proceedings of IEEE Computer Vision

Pattern Recognition'96, Computer Society Conference [C], Washington, DC, USA, 1996: 425~430.

- Chang T, Kuo C-C J. Texture analysis and classification with tree-structured wavelet transform [J]. IEEE Transactions on Image Processing, 1993, 2(10): 432~435.
- Shirazi M N, Noda H, Takao N. Texture modeling and classification in wavelet feature space [J]. IEEE Transactions on Image Processing, 2000, 1(9): 272~275.
- Porat M, Zeevi Y Y. The generalized gabor scheme of image representation in biological and machine vision [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1988, 10(7): 452~468.
- Kingsbury N G. The dual-tree complex wavelet transform; a new efficient tool for image restoration and enhancement [A]. In: Proceedings of EUSIPCO 98 [C], Rhodes, Greece, 1998: 319~322.
- Kingsbury N G. Shift invariant properties of the dual-tree complex wavelet transform [A]. In: Proceedings of International Conference America Society Signal Processing 99 [C], Phoenix, Arizona, USA, 1999: 16~19.
- 张学工. 关于统计学习理论与支持向量机 [J]. 自动化学报, 2000, 26(1): 32~43.
- 边肇祺, 张学工. 模式识别 [M]. 北京: 清华大学出版社, 2000: 286~294.
- Vapnik V N. Statistical learning theory [M]. New York: John Wiley & Sons, 1998.
- Platt J. Fast training of SVM using sequential minimal optimization advances in kernel methods support vector machine [M]. Cambridge Boston, MA, USA: MIT Press, 1999: 185~208.



王俊卿 1974 年生,1998 年获吉林大学计算机专业软件专业理学学士学位,现为中国科学院沈阳自动化研究所博士研究生,主要从事图像处理、模式识别研究。
E-mail: wangjunqing@ms.sia.ac.cn



黄莎白 1939 年生,1963 年毕业于中国科学技术大学自动化系自动控制理论专业,现为中国科学院沈阳自动化研究所研究员,博士生导师。主要从事模式识别、智能控制、计算机仿真、虚拟现实研究。



史泽林 1965 年生,1990 年获西安电子科技大学计算机应用专业工学硕士学位,现为中国科学院沈阳自动化研究所研究员,博士生导师。主要从事图像分析处理与目标识别方面的研究。