

基于小波多尺度分析的图象快速匹配模型

刘海鹰 黄胜华

(中国科学院自动化研究所, 北京 100080)

彭思龙

(中国科学院数学研究所, 北京 100080)

洪继光

(国家专用集成电路设计工程技术研究中心, 北京 100080)

摘要 提出了小波多尺度分析理论在计算机中层视觉二维图象识别中的一种应用, 并给出了一个高效匹配模型。利用二维图象的小波多尺度分解来模拟2个最主要的生物视觉通路, 并对不同通道分别进行匹配计算, 由粗到精地实现了基于扫视和凝视的二维图象的快速识别。与传统的模板匹配法相比, 该方法具有速度快、搜索范围大、精度高、抗噪能力强等优点。

关键词 图象识别, 计算机视觉, 小波变换, 多尺度分析

1 引言

根据已知图象模式从另一幅图象中寻找相同图象模式的过程叫图象匹配。图象匹配是计算机视觉的重要环节和应用之一。模板匹配法是图象匹配的典型方法。其基本思想是定义模板和搜寻视野中匹配窗口之间的距离, 在视野上逐点计算并比较该距离, 取较大值处作为匹配结果。传统模板匹配法概念清晰, 实现简单, 但需计算搜寻区域上所有点, 所以效率很低。对多数应用, 待搜寻目标在搜寻区域上往往是少数, 这意味着大量计算浪费在根本不可能是目标的点上。

本文从视觉理论出发, 以小波多尺度分析为数学工具, 建立了一套有效的中层视觉匹配模型。借助小波分解的多尺度特性及快速算法, 在小尺寸粗尺度图象上排除大量不可能是目标的点, 把主要计算集中在细尺度图象的极少数点上, 实现了扫视与凝视相结合的由粗到精的快速图象匹配。

2 二维视觉匹配模型

相当一部分视觉系统, 根据二维图象的特征就可识别物体^[1]。对于满足下列条件的应用, 都可利用二维图象建立识别算法模型^[1]:

- (1) 物体厚度与成像设备到物体间距相比可忽略;
- (2) 不考虑物体的空间深度, 且物体平面与图象平面平行或大致平行;
- (3) 物体到摄像机的投影可近似认为是平行投影。

显微图象识别、原子检测等, 都属于这类应用。神经生物学上, 人类视觉信息处理是在基本相互独立的几个通道上完成的^[1]。每种神经细胞只对某一频段的信息敏感。如图象中突出灰度分布的低频部分和突出边缘的高频部分。心理学上, 人类视觉系统对图象的识别又是一个心理感知过程, 这个过程往往是由粗到精的。小波多尺度分析恰恰为我们提供了实现这种视觉模型所需的数学工具。

* 本文是国家科技部重点科技项目:“亚微米、深亚微米集成电路自动化分析技术研究”课题之一, 并受其基金资助

收稿日期:1998-04-22;收到修改稿日期:1998-05-28

3 模型的基本思路

实际图象 $f(x, y)$ 都是有限能量的。设 $f_0 \in V_0 \in L^2(\mathbf{R}, \mathbf{R})$ ，则利用小波变换可将 f_0 分解为 4 幅图象。其中 f_j^m 表示第 j 层上的第 m 幅分解图象。对于某一层上的分解： $m=0$ 为图象的低频分量，包含图象的主要信息； $m=1$ 和 $m=2$ 分别为图象边缘水平和垂直分量，合适小波基下具有局部极大模值的向量 $f_j^2(x, y) + if_j^1(x, y)$ 是该尺度下图象的一种边缘描述^[3]； $m=3$ 为图象对角分量，通常可认为它包含了图象孤立点噪声。基于以上观点，我们构造了如图 1 的二维图象识别系统模型。

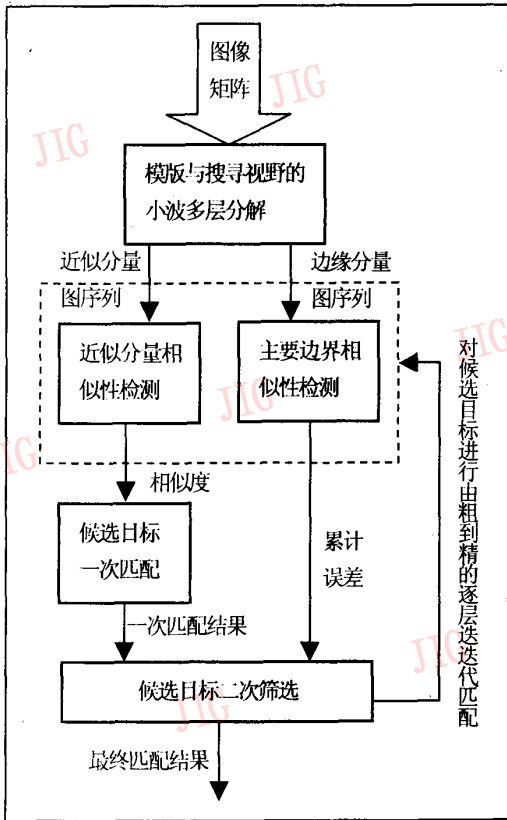


图 1 基于小波多尺度分析的二维图象快速识别系统

以保证分解后图象信息的相对位置不变。模板和视野的分解次数主要由模板大小、搜索精度和效率综合考虑决定。对于 $N \times M$ (N 行 M 列) 模板，设经过 j 次分解后一个分量图象大小为 $N_j \times M_j$ ，并设 $M'_j = \max(M_j, N_j)$ ，则 $j = \lceil \log_2(\max(M, N)/M'_j) \rceil$ 。为保证分解后图象匹配的精度和效率，通常可取 $M'_j = 8$ 。

搜寻视野的大小 $L_{行} \times L_{列}$ 可取模板大小的数倍，且 $L_{行} = 2^p, L_{列} = 2^q, p, q \in Z$ 以减少图象尺寸的奇偶性对二尺度小波分解数据精度的影响。分解 j 次后图象及数学标记如图 2 所示。

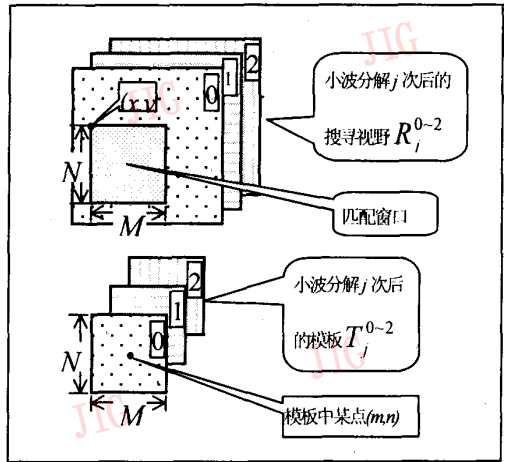


图 2 图象的小波分解参数示意图模型

4.2 图象近似分量一次匹配

图象小波分解低频分量包含了图象主要信息，是上层图象的近似，因此首先用它来定义该尺度上待搜寻窗口与模板之间的相似度。对于一类视觉应用，如显微图象识别，真实图象常被掩盖在高噪声下，若直接以绝对误差定义相似度，难以达到满意的匹配效果。为此，我们把模板图象和待搜寻窗口按相同顺序展开成向量，并定义它们之间的夹角余弦为二者的相似度。它比较的是模板与匹配窗口之间的相对相似程度，能较好地排除上述噪声影响。如图 2 所示，定义一次匹配相似度为：

$$s(x, y) = \cos(\vec{V}_{T_j^0}, \vec{V}_{R_j^0(x,y)}) = \frac{\vec{V}_{T_j^0} \cdot \vec{V}_{R_j^0(x,y)}}{\|\vec{V}_{T_j^0}\| \cdot \|\vec{V}_{R_j^0(x,y)}\|} \quad (1)$$

设经 j 次分解后搜索区域图象尺寸为 $L_{行} \times L_{列}$ ，则经上述计算后得到的相似度矩阵为 $S_{(L_{行}-M_j) \times (L_{列}-N_j)}$ 。由于 $-1 < s(x, y) < 1$ ，我们利用二分法在该尺度下进行匹配。具体算法如下：

4 识别算法设计

4.1 图象小波分解

位置信息对于识别是至关重要的，因此所选择的小波滤波器应具有线性相位，如双正交小波基等，

(1) 构造相似度局部极大值点位置集合 M_j ;

(2) 初始化 $s_{\text{下限}}, s_{\text{上限}} = 1$;

(3) $s_{\text{阈值}} = (s_{\text{下限}} + s_{\text{上限}}) / 2$;

(4) 定义候选目标集合:

$$W_{j,1} = \{(x, y) | (x, y) \in M_j, s(x, y) \geq s_{\text{阈值}}\};$$

(5) 若 $\exists (x_1, y_1), (x_2, y_2) \in W_{j,1}, |x_1 - x_2| < \epsilon_1 M_j$ 且 $|y_1 - y_2| < \epsilon_2 N_j$, 即候选目标之间严重重叠, 则 $s_{\text{下限}} = s_{\text{阈值}}$ 返回步骤(3), 式中实系数 $\epsilon_1, \epsilon_2 \in (0, 1)$ 决定了算法对漂移的容忍程度;

(6) 否则 $s_{\text{上限}} = s_{\text{阈值}}$ 并返回步骤(3), 直至 $|s_{\text{上限}} - s_{\text{下限}}| < \delta$. 其中 δ 为一很小的正数, 它决定了阈值的精度和二分法收敛的快慢。

经过上述快速叠代, 就得到了一次匹配的候选目标位置集合 $w_{j,1}$

4.3 图象主要边缘二次筛选

单纯在粗尺度上利用小波分解图象近似分量做匹配可能会虚报较多假目标。这是因为:

(1) 数学上, 完全可用该近似分量和任意 3 幅同样大小的图象通过相应的小波重构算法得到细尺度图象。这种图象尽管在视觉上可能与原图象大相径庭, 但其小波分解近似分量却可以完全相同。因此不同图象的小波分解近似分量可能相同。

(2) 视觉理论上, 图象的边缘也是识别物体的重要特征之一。因此, 在匹配中忽略图象边缘, 可能出现较多的误判。

可见, 在利用小波分解图象近似分量进行一次匹配识别之后, 对其候选目标进行主要边缘的二次筛选是非常必要的。另外在粗尺度上排除的非目标点越多, 越能减少细尺度上对候选目标所进行的匹配计算量。具体算法描述如下:

(1) 定义向量:

$$\vec{V}_{T_j}(m, n) = T_j^2(m, n) + iT_j^1(m, n),$$

$$\vec{V}_{R_j, x, y}(m, n) = R_j^2(x+m, y+n) + iR_j^1(x+m, y+n)$$

(2) 计算模板在尺度 j 下主要边界梯度集合: $E_{T_j} = \{\vec{V}_i | \vec{V}_i \in \vec{V}_{T_j}, \|\vec{V}_i\| \text{ 为局部极大值, } \|\vec{V}_i\| > E(\|\vec{V}_{T_j}\|)\}$, 记 N_{E_j} 为主边界个数; 定义该尺度下相应一次匹配候选目标的主要边界集合 E_{R_j} , 它由对应 E_{T_j} 中各向量位置处的 \vec{V}_{R_j} 构成。

(3) 对于各主要边界向量 $\vec{V}_T(i) \in E_{T_j}, \vec{V}_R(i) \in E_{R_j}$ 计算累计边界误差: $e_j = \sum_{i=1}^{N_{E_j}} \text{Ang}(\vec{V}_T(i), \vec{V}_R(i))$, 其中 $\text{Ang}(\cdot, \cdot)$ 为两向量夹角。

(4) 对每个候选目标重复步骤以上三步, 剔除主要边缘误差 e_j 较大者, 即得到 j 层上的匹配结果。

4.4 候选目标的多尺度扫视与凝视

根据小波理论, 一幅图象经小波多尺度分解后得到的图象近似分量 $f_j(x, y) \in V_j$ 是原图象的一个不同程度的近似表示, 且有 $\dots \subset V_j \subset V_{j-1} \subset \dots \subset V_0$ 。这意味着:

(1) 相同成像条件下得到的搜索视野图象和模板图象同属原始空间 V_0 , 因此两图象使用合适小波基经过 j 次小波分解后, 其近似分量仍属同一空间 V_j , 它们之间可用同样的相似度定义进行相似性检测。从而保证了真实目标被包含在粗尺度搜索视野内的候选目标中;

(2) 由于每次做进一步小波分解时抛弃了高频分量, 因此完全可能出现小波分解的近似分量相同而原图象完全不同的情况。尽管在第 j 层的匹配中利用该层的 2 个高频分量对候选目标进行过筛选, 但这 2 个高频分量不包含第 $j-2$ 层及以上层次的高频信息。所以粗尺度图象的候选目标中可能包含假目标。

(3) 随图象尺度变粗, 图象位置信息也一步步丧失, 粗尺度图象上的目标无法直接精确定位在原图中。

因此在第 j 层上对图象进行搜索匹配运算后, 需在第 $j-1, j-2, \dots$ 层中针对上一层候选目标逐层进行更精细的匹配。随着图象内容和位置信息越来越完整, 假目标将被层层排除, 直至原图, 从而得到最终目标及其精确的定位。由于这时每层上的搜索匹配运算只需在上层候选目标的对应位置上进行, 所以匹配数据大大减小, 匹配效率大大提高。

4.5 计算量估计

整个算法的计算量主要有图象小波分解、相似度计算、和特征边缘二次筛选 3 个部分。

下面, 我们分析占总体计算量中比例最大的两个部分: 小波分解的计算量和相似度的计算量, 并着重讨论它相对于传统模版匹配法所节约的计算量。讨论中设一次乘法的计算量为 1, 搜索范围图象的大小为 $M \times N$, 模版图象的大小为 $m \times n$ 。

设小波分解滤波器长度分别为 $l_{\text{low}}, l_{\text{hi}}$, 则完全小波分解的标准快速算法的计算量为

$$Q_0 = (l_{\text{low}} + l_{\text{hi}})(MN + mn) \quad (2)$$

我们的算法模型, 选用二次样条小波, 其滤波器

的长度为 $l_{low}=5, l_{hi}=3$ 。

为了使方法之间具可比性,设我们的算法模型和传统模板匹配法采用相同的相似度定义,且定义一次相似度运算的等价乘法计算量为 k 。则:对于传统匹配法,总计算量为

$$Q = k(M - m)(N - n)mn \quad (3)$$

对本文提出的算法,设图象总共分解了 j 次,则在第 j 层上的计算量为

$$Q_1 = k(M - m)(N - n)mn/16^j \quad (4)$$

它仅为式 (3) 的 16^j 分之一。

设在第 j 层上的候选目标个数为 p ,计算量最大的情况是这个 p 候选目标全部是最终的真实目标,再考虑小波二尺度分解中粗尺度图象上的一个像素点对应相邻细尺度上的 4 个像素点,则:第 $j-1$ 层至原图象上相似度的计算量之和为

$$Q_2 = (16/3)kpmn[1 - (1/4^j)] \quad (5)$$

则总的计算量约为

$$Q' = Q_0 + Q_1 + Q_2 \quad (6)$$

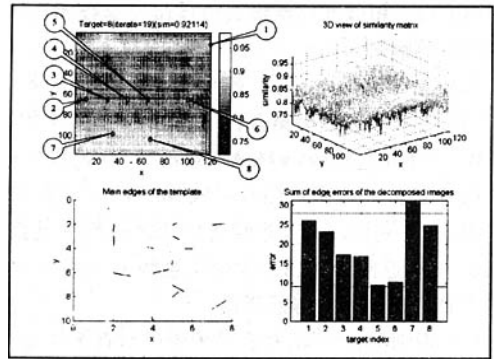
对于式(1)相似度定义: $k \approx 4$ 。当取集成电路版图分析中的典型值: $M=N=1\ 024, m=n=100, j=3$ 时, $Q \approx 3.42 \times 10^{10}, Q' \approx 1.68 \times 10^7 + 2.1 \times 10^5 p$ 。这时,即便 $p=100$ (这已经是非常极端的情况),总计算量仍然比传统模版匹配法的计算量减少约 2~3 个数量级。而且图象搜索区域越大,节省的计算量越多。

5 实验结果

以下是本方法在集成电路显微版图分析中的应用。所有算法均在 MATLAB 5.0 上实现。图 3 是在尺度 3 上的匹配。图 3 (a)、(b)是一次粗匹配相似度矩阵的侧视图,计算中取最小相似度为 0.7, $\epsilon_1 = \epsilon_2 = 0.8, \delta$ 为最大相似度与最小相似度差值的千分之一。图 3 (a)中亮点和图 3(b)中尖峰是相似度值较大处。图 3(c)是该尺度上的主要边界向量。图 3(d)是候选目标主要边缘累计误差,实验中排除累计误差超过最小值 3 倍的假目标(目标⑦)。

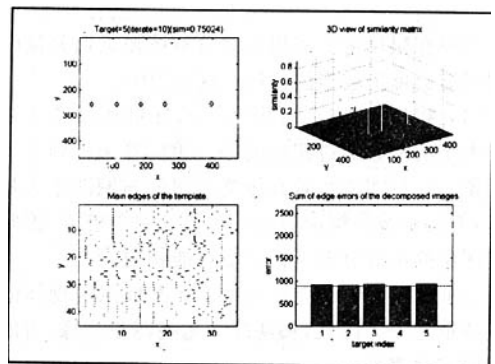
做为对照,图 4 给出了 $j=1$ 层上的匹配结果。

图 5 是最终匹配结果。整个匹配过程历时约 2 分钟,相同计算环境下使用 FFT 做直接相关法历时大于 3 小时,该方法计算速度较 FFT 做直接相关法提高了约 2~3 个数量级,与理论估计相符,而且匹配结果大大优于直接相关法。



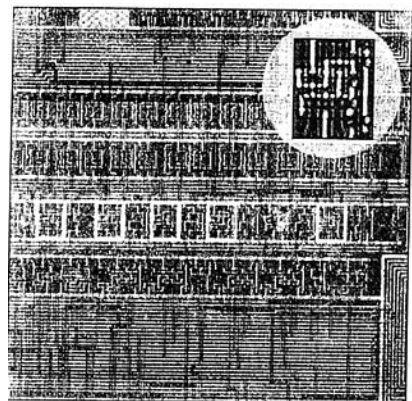
$j=3$ 层上的匹配;(a) 相似度矩阵平面图,菱形框处是一次匹配候选目标位置;(b)相似度侧视图;(c) 主要边界向量;(d) 候选目标主要边缘误差累计(图中标明了最小误差及其 3 倍处)

图 3



$j=1$ 层上的匹配;(a) 相似度矩阵;(b) 相似度侧视图;(c) 主要边界向量;(d) 候选目标主要边缘误差累计

图 4



原图最终匹配结果,方框处标识出了目标所在位置(右上角为单元模板图象)

图 5

6 结论

本文基于小波多尺度分析提出了一个二维图象的扫视与凝视的匹配识别模型。该模型不是直接对原尺寸大小的搜索区域进行匹配,而是先利用小波多尺度分析将图象分解到粗尺度空间并对其进行扫视,在小尺寸粗尺度图象上排除大量不可能出现目标的位置后,对剩下的少量候选目标进行更精一层尺度上的凝视匹配,由粗到精地进行快速识别,从而识别出模板目标。理论和实践证明,该方法具有匹配速度快、精度高、抗干扰能力强等特点,尤其适合于大范围搜索视野下目标的识别。

参考文献

1 马颂德,张正友. 计算机视觉—计算理论与算法基础. 北京: 科学

出版社,1998.

2 Mallat S, Zhong Sifen. Characterization of Signals from Mutiscale Edges. *IEEE Trans. on PAMI*, 1992, 14(7).

3 Barnea D I, et al. A Class of Algorithm for Digital Image Registration. *IEEE Trans. Comput.* C-21, 179~186.

4 Ingrid Daubechies, Ten Lectures on Wavelets. Society for Industrial and Applied Mathematics, 1992.

5 David Marr. *Vision*, W. H. Freeman and Company, 1982.

6 钟声,石青云,程民德. 基于小波变换的立体视觉匹配方法. 模式识别与人工智能,1994,7(1).

7 Chin R T, Dyer C R. Model-based recognition in robot vision. *ACM Computer Surveys*, 1986, 18(1):67~108.

8 Hu M K. Visual Pattern Recognition by Moment Invariants. *IEEE Transaction Information Theory*, 1962, 8(2): 179~187.

9 Rosenfeld A, Kak A. *Digital Picture Processing*. New York: Academic, 1982, 2nd edit.



刘海鹰 1993年福州大学电气工程系本科毕业。现为中科院自动化所博士研究生。研究方向为:模式识别、计算机视觉、集成电路版图图象分析。



黄胜华 1993年北京师范大学数学系本科毕业。现为中科院自动化所博士研究生。研究方向为:模式识别、集成电路逻辑综合、图象处理。



彭思龙 1993年安徽大学数学系本科毕业。现为中科院数学所博士研究生。研究方向为:小波理论及其应用,侧重于数值计算、模式识别。



洪继光 1962年毕业于成都电讯工程学院。研究员,博士生导师,中科院自动化所总工,国家专用集成电路设计工程技术研究中心主任。主要研究方向为集成电路、图象处理、模式识别和计算机视觉等。

Fast Image Matching Model Based on Wavelet Multiscale Analysis

Liu Haiying, Huang Shenghua

(Institute of Automation, CAS)

Peng Silong

(Institute of Mathematics, CAS)

Hong Jiguang

(National ASIC Design Engineering Center)

Abstract In this paper, an effective matching model for intermediate level computer vision is proposed based on wavelet multi-scale analysis. Two main vision paths are simulated by wavelet multi-scale decomposition of images, and scanning and fixing algorithms are designed. With complete the fast recognition from coarse to fine scale. Compared with conventional methods, the model is faster, more accurate, and higher noise-robust.

Keywords Image recognition, Computer vision, Wavelet, Multiscale analysis