



[12] 发明专利申请公开说明书

[21] 申请号 200410035084.9

[43] 公开日 2005 年 11 月 2 日

[11] 公开号 CN 1691054A

[22] 申请日 2004.4.23

[21] 申请号 200410035084.9

[71] 申请人 中国科学院自动化研究所

地址 100080 北京市海淀区中关村东路 95 号

[72] 发明人 谭铁牛 胡卫明 杨金锋 王 谦

[74] 专利代理机构 中科专利商标代理有限责任公司

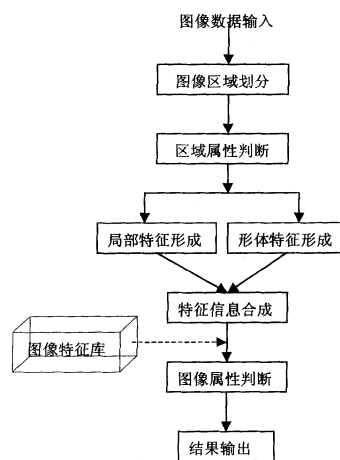
代理人 戎志敏

权利要求书 2 页 说明书 9 页 附图 3 页

[54] 发明名称 基于内容的图像识别方法

[57] 摘要

一种基于内容的图像识别方法，包括步骤：首先对图像进行网格划分；然后以网格各节点为初始位置进行区域生长来获取图像中各节点周围的肤色信息；利用网格单元区域和各节点的相互投票的机制确定图像中的兴趣点和兴趣区域；在兴趣区域和兴趣点的基础上，利用点的运动提取图像中人体躯干的轮廓；最后，提取轮廓信息和轮廓内部中的局部信息生成特征向量对图像进行识别和性质判断。本发明突破了诸如色彩直方图匹配、小波变换轮廓匹配、肤色纹理描述、图像中心矩匹配等现有敏感图像识别技术速度慢、效率低、设备依赖性强等方面的难题，解决了三点式泳装图像，裸体图像和人脸图像的分类，使敏感图像识别技术得到了进一步的扩展并开拓了广阔的应用前景。



1. 一种基于内容的图像识别方法，包括步骤：
 - 首先对图像进行网格划分；
 - 5 然后以网格各节点为初始位置进行区域生长来获取图像中各节点周围的肤色信息；
 - 利用网格单元区域和各节点的相互投票的机制确定图像中的兴趣点和兴趣区域；
 - 在兴趣区域和兴趣点的基础上，利用点的运动提取图像中人体躯干的轮廓；
 - 10 最后，提取轮廓信息和轮廓内部中的局部信息生成特征向量对图像进行识别和性质判断。
2. 根据权利要求 1 所述的方法，其特征在于所述区域生长包括步骤：
 - 确定功能块；
 - 15 每个功能块被分为 4 个子块。
3. 根据权利要求 2 所述的方法，其特征在于所述功能块的不同生长方向有不同的权重。
4. 根据权利要求 1 所述的方法，其特征在于所述相互投票的机制包括：
 - 对于落在图像边界上的点均赋予 0 值；
 - 20 对于落在图像区域内的满足一定肤色面积比例阈值的点均赋予 1 值，否则赋予 0 值；
 - 区域的最终得分值由它周围的各点和它本身的携带的值共同确定。
5. 根据权利要求 1 所述的方法，其特征在于所述提取图像中人体躯干的轮廓包括：
 - 25 检测肤色边界点，把一些权重更高的感兴趣点连接起来形成一条闭合曲线；
 - 采集落在所述闭合曲线外面但离曲线一定距离内的肤色边界点，将这些点连接起来形成另一条闭合曲线；
 - 调整另一条曲线上所有的非肤色边界点的位置，得到优化后的躯干轮廓曲线；
 - 30

检测轮廓内的非肤色区域，得到局部信息。

6. 按权利要求 5 所述的方法，其特征在于所述局部信息包括非肤色区域的面积和位置。

7. 按权利要求 1 所述的方法，其特征在于所述对图像进行识别和性质判断包括提取如下特征：

感兴趣区域的位置几何信息；

感兴趣区域的得分及最高分的位置；

轮廓内相邻的非肤色区域之间的夹角；

轮廓内面积最大的非肤色区域的位置。

基于内容的图像识别方法

5 技术领域

本发明涉及模式识别领域，特别涉及基于内容的图像识别方法。

背景技术

随着现代互联网技术的飞速发展，网络对全球经济和社会生活的渗透，其深度和影响力已远远超出人们的预料。网络信息安全逐渐成为一个极为重要的问题，其中对社会，尤其是对未成年人的影响更引起人们的广泛关注，所以信息过滤技术成了迫切的理论和实际需求。在美国，这些问题早在1994年就已引起公众的注意，当时美国社会深受便利的网络在线色情图片、暴力、恶毒言论等有害网络信息的痛苦困扰，诸多新闻、报纸、杂志都充斥着对色情网站、各种丑恶团体、在线性侵犯等问题的恐惧。网络上大量的有害内容直接导致国会通过“Communications Decency Act(CDA)”和“Child On-line Protection Act(COPA)”两项法律。以此为法律依据，美国的软件工业开发了他们自己的网络内容过滤软件(Content blocking filter software)并建立网络内容审计系统平台(Platform for Internet Content Selection-PICS)。随后在1999年国会又通过了“Children’s Internet Protection Act(CIPA)”法律以保护青少年，使其免受网络有害信息的侵害。

在对敏感信息检测方面，国外一些大学(Berkeley, Iowa, Standford)开展了部分对网络上敏感图片进行分析的探索。Fleck 与 Forsyth[D. A. Forsyth, M. M. Fleck, Body plan, Proc. IEEE Conference on Computer Vision and Pattern Recognition, 1997, pp.678-863.]通过检测人体的皮肤，并把各部分皮肤区域连成一组，来识别一幅图片是否包含裸体内容。James Ze Wang [J. Z. Wang, G. Wiederhold, O. Firschein, System for screening objectionable images, Computer Communications Journal, Elsevier Science, 1998, 21(15), pp.1355-1360.]利用 WIPE

(Wavelet Image Pornography Elimination) 方法对敏感图片进行识别与过滤。这种方法综合利用了 Daubechies 小波变换, 归一化中心矩, 颜色直方图形成语义匹配向量进行图像分类识别。Jones 和 Rehg [M. J. Jones, J. M. Rehg, Statistical color models with application to skin
5 detection, Proc. the International Conference on Computer Vision and Pattern Recognition, 1999, pp.274-280.]对统计肤色模型作了深入的研究, 他们先从互联网上采集了大量图像, 并且手工标注了图像中的肤色区域, 然后把这些图像作为训练样本从中训练中肤色模型, 最后以肤色为主要信息去检测敏感图像。Bosson 等人 [A. Bosson, G. C.
10 Cawley, Y. Chan, R. Harvey, Non-retrieval: Blocking pornographic images, Proc. the International Conference on Image and Video Retrieval, 2002, pp.50-60.]把检测出的肤色区域作小椭圆块处理, 对每个椭圆都提取诸如面积, 中心点, 轴长, 离心率等特征, 并对这些特征进行分类。此外, 还有一些通用的基于内容的图像检索系统, 如 IBM
15 的 QBIC, Attrasoftware 的 ImageFinder、MW Labs 的 Imatch 等。值得一提的是, 法国国立计算机与自动化研究所(INRIA)图像与多媒体索引组的四名科学家于 1999 年成立了 LookThatUp 公司, 该公司图像过滤与检索产品在行业内处于领先地位。LookThatUp 的 Image-Filter 利用先进的识别算法可以对网络上的图像进行实时过滤。

20 在 2001 年, 欧洲启动 NetProtect 计划, 该计划从 2001 年 1 月 1 日开始到 2002 年 5 月 1 日结束, 由法国的 EADS Matra Systemes & Information 研究机构联合西班牙的 Red Educativa、法国的 Matra Global
Netservices、希腊的 Hyertech、德国的 SailLabs 等科研机构共同开发研制。NetProtect 计划的目标是建立欧洲互联网内容过滤工具的统一技术
25 标准, 以实现跨地域、跨语言的互联网有害信息过滤。

国内现有的反黄软件有美萍软件工作室推出美萍反黄专家、ZiJinCheng.NET 推出的紫禁城反黄卫士、飞涛软件工作室推出的护花使者 (像爱护花朵一样保护青少年, 使其不受电子黄毒的侵害)、中国科学技术大学讯飞信息科技有限公司开发的“火眼金睛”电脑反黄软件、清华
30 华的“五行卫士”电脑反黄软件等等。需要提出的是国内的这些网络有害

信息过滤软件无论从技术上还是从过滤方法上均不能达到应有的效果。特别是我国近些年来网络应用的飞速发展，致使网络对社会、家庭、教育的影响日益深远，所以网络有害信息过滤将面临前所未有的压力。

需要强调的是，虽然互联网有害信息过滤技术在世界范围内得到了广泛的关注和研究，但在对有害信息识别技术方面仍然存在诸多难点，其中基于图像内容的色情图片识别和过滤方法仍然缺乏有效的算法和分类方法。因此如何开发更加鲁棒、准确的敏感图像识别技术仍是一个挑战。

发明内容

10 本发明的目的是提供一种基于内容的图像识别方法，解决的技术问题是利用敏感图片所能表达的人体的局部和形体信息来达到对敏感图像的认识。

为实现上述目的，一种基于内容的图像识别方法，包括步骤：

首先对图像进行网格划分；

15 然后以网格各节点为初始位置进行区域生长来获取图像中各节点周围的肤色信息；

利用网格单元区域和各节点的相互投票的机制确定图像中的兴趣点和兴趣区域；

20 在兴趣区域和兴趣点的基础上，利用点的运动提取图像中人体躯干的轮廓；

最后，提取轮廓信息和轮廓内部中的局部信息生成特征向量对图像进行识别和性质判断。

25 本发明是一种新型的敏感图像识别技术，突破了诸如色彩直方图匹配、小波变换轮廓匹配、肤色纹理描述、图像中心矩匹配等国际现有敏感图像识别技术速度慢、效率低、设备依赖性强等方面的难题，同时解决了三点式泳装图像，裸体图像和人脸图像的分类难题，使敏感图像识别技术得到了进一步的扩展并开拓了广阔的应用前景。

附图说明

30 图1是图像的几何划分,其中,图(a)是区域和点, 图(b)是区域和周围子

区域的关系，图(c)是点和周围子区域的关系；

图2是区域生长，其中，图(a)是4个单位向量，图(b)是生长方向由合成向量决定；

图3是区域投票，其中，图(a)是投票结果，图(b)感兴趣的区域，
5 黑色部分为非目标区域；

图4是提取躯干轮廓和局部信息的几个过程；

图5是初始曲线和参考项，其中，图(a)是由感兴趣点组成的初始曲线，图(b)是参考向量和参考点，图(c)是点的运动模式；

图6是图像分类形式的拓扑结构；

10 图7是图像识别与判断流程。

具体实施方式

本发明的主要特点在于：1) 采取了一种新型的图像网格划分和区域生长技术。这种技术能快速有效的提取图像中的肤色信息；2) 在网络
15 单元和节点投票的基础上，获取兴趣点和兴趣区域，这种方式缩短了目标区域获取的时间，降低了计算代价；3) 综合利用兴趣区域和兴趣点作为人体躯干轮廓提取的初始信息，通过像素点集的优化来生成躯干轮廓。这个过程不仅能获取图像的局部信息，同时也能表达人体的形体信息；4) 提取人体的形体特征、轮廓特征、局部特征并确立图像快速分类的方法。

20 下面详细给出该发明技术方案中所涉及各个细节问题的说明：

图像网格划分

对图像进行适当的划分是图像处理中经常采用的方法，这里我们对图像进行划分的目的是为了节约对图像进行底层处理的时间，同时也是为了方便地定位图像的感兴趣区域。划分方法如附图 1(a)所示，图像被
25 分为 4×4 共 16 个相等的区域。每个区域表示为 $a_{i,j}$ ，其中 $i, j=1,2,3,4$ 。在每个区域的 4 个角上分别还标出 4 个子区域，因此对于一幅图像而言，每个区域和它所对应的 4 个子区域以及角落上的点之间都有联系，角落上每个点将周围 4 个区域联系在一起。我们对这些区域、子区域和点分别进行编号，如附图 1(a), (b)和(c)所示。根据图 7(b)和(c)，我们可以定
30 义下面 2 个矩阵

$$A = \begin{bmatrix} p_{ij}^{(4)} & p_{i,j+1}^{(3)} \\ p_{i+1,j}^{(2)} & p_{i+1,j+1}^{(1)} \end{bmatrix} \quad P = \begin{bmatrix} p_{ij}^{(1)} & p_{ij}^{(2)} \\ p_{ij}^{(3)} & p_{ij}^{(4)} \end{bmatrix} \quad (1)$$

A 表示区域和它周围的子区域之间的关系。 P 表示角落点和它周围的子区域之间的关系。然后就可以用区域、点和子区域来表示整幅图像。这将大大节省计算复杂度，为下面的步骤奠定基础。

5 区域生长

考虑到网络图像往往具有不同的尺寸，我们采用了一种新的区域生长方法，它以功能块而不是像素点为基本单位进行生长并利用已经得到的肤色分布来检测子区域内的肤色信息。首先我们从点 p_{ij} 出发，以它为中心在其周围确定一个 6×6 的功能块。然后确定 4 个单位向量，他们分别从中心点 p_{ij} 指向块的 4 个角，如附图 2(a)所示。这 4 个向量的叠加可进一步合成新的向量，用于确定块的生长方向，如附图 2 (b)所示。不同方向有不同权重，为确定这些权值，一个块被均分为 4 个子块，对每个子块都计算其肤色面积比，并选择该比值作为对应的单位向量的权。例如，

$$\vec{s}_g = w_i \vec{s}_i + w_j \vec{s}_j \quad (2)$$

其中 w 为权值。整个块生长过程直到遇到子区域边界或 4 个权值都退化为 0 时才结束。我们定义 S_{grow} 为生长区域总面积， S_{skin} 为肤色面积。实验表明，虽然 S_{grow} 包含了生长区域内的非肤色面积，但在描述区域和点的属性时， S_{grow} 比 S_{skin} 更有效。于是我们定义

$$p_{ij}^{(e)} = S_{grow} / S_{sub} \quad (e = 1, 2, 3, 4) \quad (3)$$

20 兴趣点和兴趣区域的生成

不可否认图像中一些不具有肤色信息的区域属于敏感图像识别中的冗余信息，有效排除这些冗余区域将会缩减提取图像有用信息的时间。直观上讲，肤色丰富的区域应该是算法最感兴趣的区域，所以在这里我们采用了基于肤色信息的投票机制来获取感兴趣区域。我们定义 a_{ij} 的取值为

$$a_{ij} = \sum (p_{ij}^{(4)} + p_{i,j+1}^{(3)} + p_{i+1,j}^{(2)} + p_{i+1,j+1}^{(1)}) \quad (4)$$

这里方程右端的各个分量由方程(3)给出。然后对这些表达区域属性的变量进行归一化得到

$$a_{ij} = a_{ij}/M \quad (5)$$

其中 $M = \max(a_{11}, a_{12}, \dots, a_{ij}, \dots, a_{44})$, $M \neq 0$ 。为突出区域之间的差别, 还须确定点 p_{ij} 的可能取值。我们定义 p_{ij} 的值为

$$p_{ij} = \begin{cases} 1 \text{ (true),} & \text{if } \sum_{e=1}^4 p_{ij}^{(e)} \geq \alpha \\ 0 \text{ (false),} & \text{if } \sum_{e=1}^4 p_{ij}^{(e)} < \alpha \end{cases} \quad (6)$$

- 5 其中 α 是一个阈值。这里对于落在图像边界上的点均赋予 0 值, 因为从拍摄角度来看, 所要描述的物体应该位于靠图像中间的位置。选择那些取值为 1 的点为感兴趣点, 则感兴趣点的总数为

$$N = \sum_{i=2, j=2}^{i=5, j=5} p_{ij} \quad (7)$$

- 10 考虑到点和区域之间的关系, 我们可以利用感兴趣点的取值来进行投票, 也就是说每个感兴趣点都将为它的四周区域投 1 票。如果图像中所有的内部点都为真, 那么投票结果如附图 3(a): 每个区域的最终票数都取决于位于它角落上的 4 个点

$$v_{score}(ij) = p_{ij} + p_{i,j+1} + p_{i+1,j} + p_{i+1,j+1} \quad (8)$$

区域的最终得分为

$$15 \quad S_{region}(i, j) = a_{ij} + v_{score}(ij) \quad (9)$$

- 我们把区域按照其得分进行降序排列, 选出从 1 到 N 共 N 个感兴趣区域并排除其它区域, 那么目标区域就能凸现出来。一些结果如附图 3(b) 所示。感兴趣区域可用于表达人的形体信息, 得分更高的区域可能包含更重要的信息和更多的可疑内容。重点研究感兴趣区域可以进一步描述形体, 并可降低计算复杂度。

提取躯干轮廓和局部信息

- 25 对于敏感图像属性的判断, 我们认为躯干包含的信息最能描述图像的性质。所以提取物体轮廓和轮廓内的局部信息是这部分的核心内容。附图 4(a-d) 给出了轮廓提取的过程。首先, 我们设计了一个肤色边缘检测器来检测肤色边界点, 然后把一些权重更高的感兴趣点连接起来形成一条闭合曲线, 如附图 4(a)所示。接下来, 采集落在该闭合曲线外面但是离

曲线在一定距离内的肤色边界点，将这些点连接起来形成另一条闭合曲线，如附图 4(b)图所示。然后再调整该曲线上所有的非肤色边界点的位置，得到优化后的曲线即躯干轮廓曲线，如附图 4(c)所示。最后，我们通过使用点生长技术来检测轮廓内的非肤色区域，得到局部信息，如非肤色区域的面积、位置等等，如附图 4(d)所示。这里我们选了一幅包含 4 个感兴趣点的图像来具体描述我们的算法，如附图 5 所示。

首先定义所有肤色边界点的集合为 Q ，显然集合中有一定数量的点位于轮廓边界上，因而我们可以从集合 Q 中得到有关轮廓的初始信息。通过连接相邻的感兴趣点可得到闭合曲线 C_r ，如附图 5(a)所示。

$$C_r = l'_1 + l'_2 + l'_3 + l'_4 \quad (10)$$

在曲线 C_r 的外侧作 C_r 的平行曲线 C_{r1} ，与 C_r 之间相距 λ ， λ 是一个阈值。我们定义夹在这两条曲线之间的肤色边界点集合为 Q ，显然 Q 是 Q 的一个子集，其大小取决于阈值 λ 。把集合 Q 包含的所有点和感兴趣点互相连接起来，得到一条可以大体描述轮廓的曲线。我们定义这条曲线上所有的点集为 D ，然后定义 D 与 Q 的差集 $E = D - Q$ 。这样集合 E 同时包含了轮廓内的肤色点和轮廓外的非肤色点。接下来我们需要调整那些不在轮廓上的点的位置，使曲线更接近于真实的躯干轮廓。

在调整点的位置时，我们定义了几个基本的参考项。首先选择连接相邻感兴趣点的 4 条线段的中点作为参考点，从每个感兴趣点引出 2 条矢量指向相邻参考点，我们称之为参考矢量，如图 5(b) 所示。参考点决定了带调整点所属的区域，而参考矢量决定其运动模式。这些参考项对于我们的算法有重要意义，因为每个带调整点在初始阶段的运动都取决于其对应的参考矢量，要么靠近它，要么远离它。

接下来，我们通过图 5(c) 举例说明待调整点的具体运动模式。对于待调整的肤色点和非肤色点我们采用不同的搜索方案。假设 p_0 表示肤色点， p'_0 表示非肤色点，它们的运动模式分别表示如下：

对于肤色点 p_0 ，

$$\begin{cases} s_n = r_n + (\|s_0\| \cos \theta_0 / \|R\|) R \\ r_n = \|s_0\| \cos \theta_0 \tan(\theta_0 + n\delta_\theta) / \|r_0\| r \\ \theta_0 + n\delta_\theta \leq \varphi \end{cases} \quad (11)$$

其中

$$\begin{cases} \theta_0 = \arccos([s_0, R]/\|s_0\| \cdot \|R\|) \\ r_0 = s_0 - (\|s_0\| \cos \theta_0 / \|R\|) R \end{cases} \quad (12)$$

对于非肤色点 p'_0 ,

$$\begin{cases} s'_n = -\|s'_0\| (i \cos(\varphi + \theta'_0 - n\delta_{\theta'}) \\ \quad + j \sin(\varphi + \theta'_0 - n\delta_{\theta'})) \\ \theta_0 + n\delta_{\theta} \leq \varphi \end{cases} \quad (13)$$

5 其中

$$\begin{cases} s'_0 = -\|s'_0\| (i \cos(\varphi + \theta'_0) + j \sin(\varphi + \theta'_0)) \\ \theta'_0 = \arccos([s'_0, R]/\|s'_0\| \cdot \|R\|) \end{cases} \quad (14)$$

从中可以看出, 肤色点沿着直线运动, 不但可以很快地检测到邻近的肤色边界点, 还可以定位距离比较远的肤色边界。而非肤色点沿着圆周运动, 可以更好地符合躯干的弯曲状况。虽然肤色点和非肤色点具有不同的运动模式, 它们的目的是是一致的, 都是为了检测落在初始轮廓外的新肤色边界点, 从而在某种程度上去逼近真实的躯干轮廓。目标函数定义如下

$$F(p_n) = g_1(p_n) + g_2(p_n) + g_3(p_n) \quad (15)$$

15 方程右边第 1 项表示是否为肤色边缘点, 第 2 项表示是否为边缘点, 其中边缘由 sobel 边缘检测器得到, 第 3 项表示像素是否为肤色点。定义

$$f_1 = F(p_{n+1}) - F(p_n) \quad (16)$$

$$f_2 = F(p_n) - F(p_{n-1}) \quad (17)$$

20 如果 $f_1 \neq 0$ 且 $f_2 \neq 0$, 那么选择 p_{n-1} 、 p_n 和 p_{n+1} 3 个点中目标函数值最大的点为目标轮廓点, 以取代其对应的集合 E 中的初始点。最终优化后的曲线如图 4(d)所示, 我们称之为躯干轮廓, 许多有关人体的重要特征都包含在轮廓之内。

25 接下来我们可以在轮廓内提取局部信息。首先不难得到轮廓的中轴线, 然后我们利用点生长方法来在中轴线两边检测非肤色区域, 从而可以得到这些非肤色区域的面积和位置信息。我们称这些居于轮廓内部的非肤色信息为局部信息。

图像特征提取与性质判断

根据 4.3 节讨论, 我们事实上丢弃了图像 4 个角上的区域所含信息, 这样做是合理的, 因为图像的重点信息通常描述在图像的内部。由此不

难看出一幅图中最多有 9 个感兴趣点。根据感兴趣点的数目，我们把所有图像分为 9 大类，然后根据图像的长宽比再把每大类分为 3 子类。因而，所有图像共分为 27 类，如附图 6 所示。这 27 类总共可以区分

$$N_c = 3 \times (C_{12}^0 + C_{12}^1 + \dots + C_{12}^8 + C_{12}^9) = 12051 \quad (18)$$

5 种图像，因而这些类别足够描述敏感图像的各种不同的差别。对于每幅图像，我们提取如下特征，首先是感兴趣区域的几何信息，然后是感兴趣区域的得分以及最高分区域的位置，第 3 是轮廓内相邻的非肤色区域之间的夹角，接下来是轮廓内面积最大的非肤色区域的位置，这由区域中心点到中轴线两端的距离之比决定。把这些特征展开重新排列成如下
10 一维向量

$$v = [w_0, w_1, \dots, w_i, \dots, w_n]^T \quad (19)$$

在这里轮廓信息不用作分类特征，因为正常的人脸图像也可能具有与敏感图像相似的轮廓信息。另外，轮廓外的非肤色区域面积也不用作分类特征，这是因为在判断敏感图像时，这一特征并不稳定。

15 对于图像的认识和匹配，我们采用最近邻方法，余弦相似性度量描述如下

$$\begin{aligned} g(v, v_i) &= \arg \min_{v_i \in C_i} d(v, v_i) \\ d(v, v_i) &= 1 - \frac{v^T v_i}{\|v\| \|v_i\|} \end{aligned} \quad (20)$$

二值分类函数为

$$G(v) = g(v, v_+) - g(v, v_-) \quad (21)$$

20 其中 v_+ 和 v_- 各代表正负模板。在识别过程中，我们首先判断待测图像
的类别，然后提取其特征，将它与该类训练集中的正负样本在特征空间
内进行匹配比较并采用式 (21) 进行判断。因为在进行特征比较之前，
已经预先将图像进行了分类，所以减少了比较次数，降低了计算复杂度。
整个判断过程见附图 7。

25

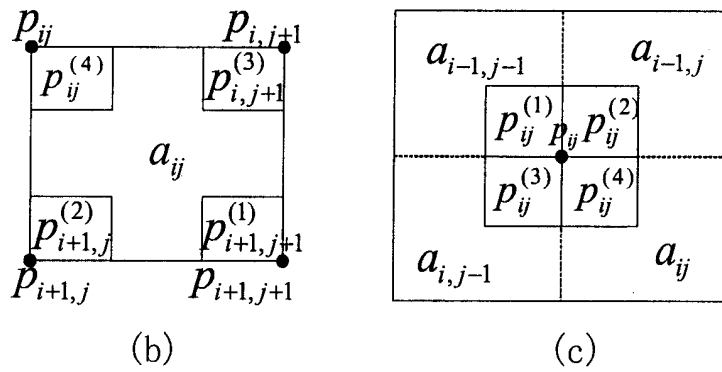
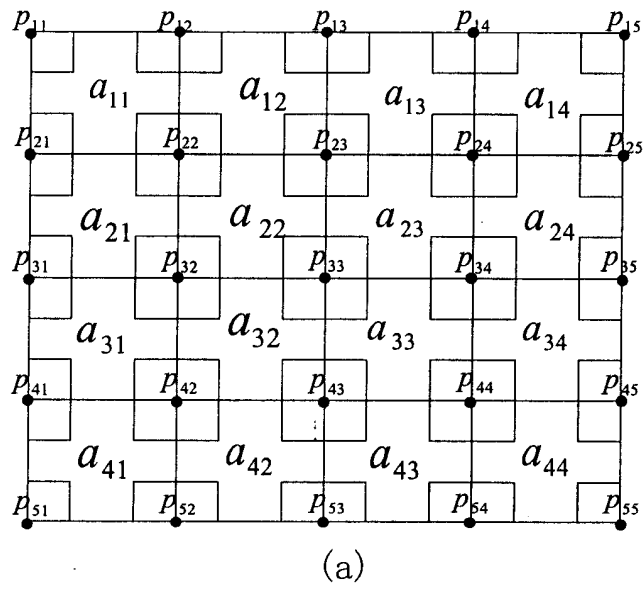


图 1

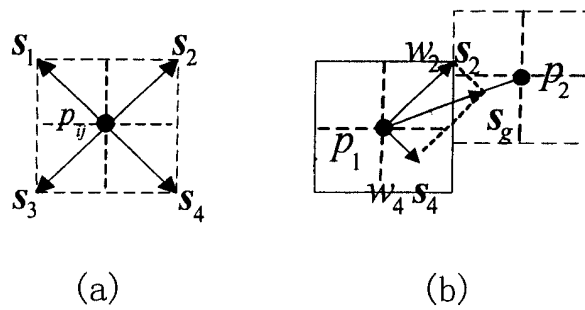
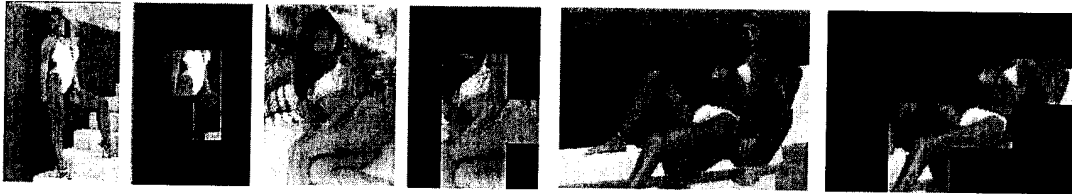


图 2

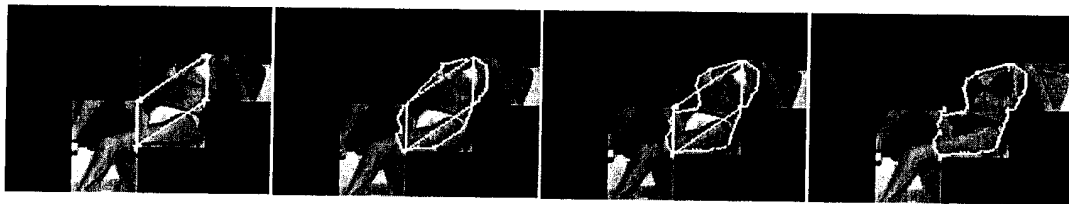
1	2	2	1
2	4	4	2
2	4	4	2
1	2	2	1

(a)



(b)

图 3



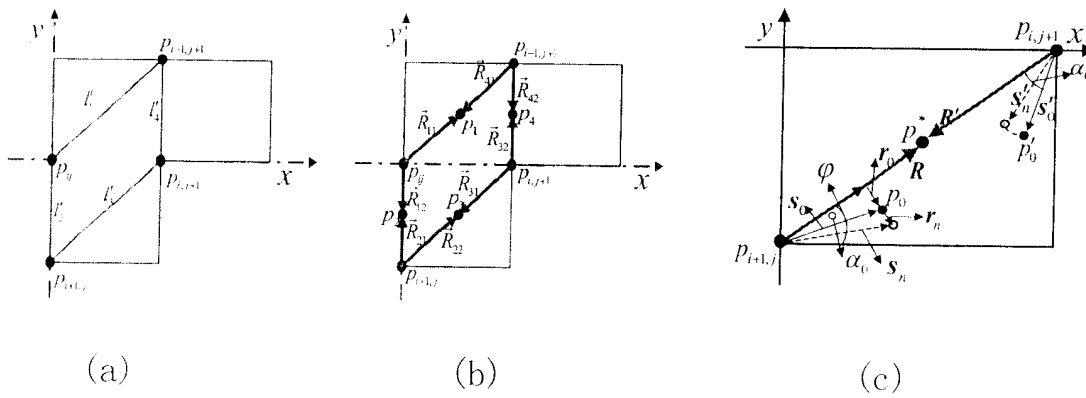
(a)

(b)

(c)

(d)

图 4



(a)

(b)

(c)

图5

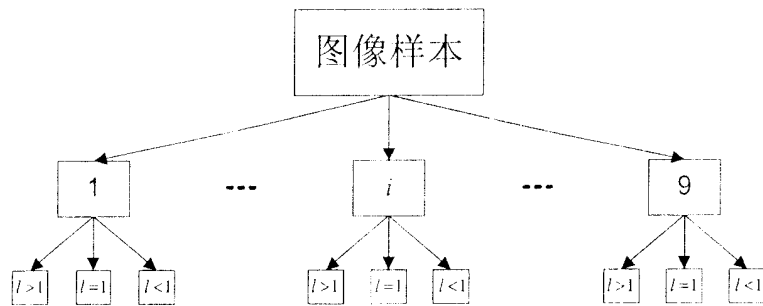


图 6

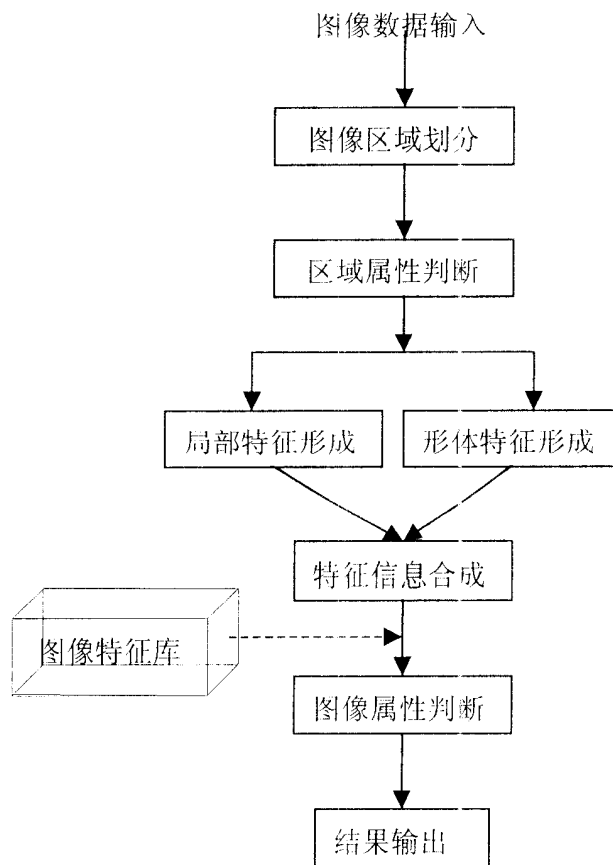


图 7