

基于小波变换的多尺度立体视觉方法*

钟 声 石青云 程民德

(北京大学数学系, 北京 100871)

(北京大学信息科学中心, 北京 100871)

(北京大学视觉听觉信息处理国家重点实验室, 北京 100871)

摘要 本文给出了一个基于二进小波变换的多尺度双眼视觉匹配算法. 对原立体图对进行二进小波变换, 得到原图对的一个多频道表示; 对变换系数的分析使我们得到了有效的消除噪声, 再提取出边缘的方法, 提出的边缘作为对应点匹配的基本特征是恰当的. 匹配的一些基本约束, 如边缘方向大致一致等, 已经包括在小波变换的系数中. 更为重要的是我们在小波变换系数上推导出4个有效的新的约束——平滑分量梯度约束和平滑分量Laplace约束, 由粗及精的多尺度控制策略得以自然引入, 使得引起误匹配的可能对应点大量减少, 从而提高了匹配的精确度和速度, 鲁棒性增强. 本方法是解决立体视觉匹配问题的更为有效的模型.

关键词 立体视觉, 对应点匹配问题, 二进小波变换, 视差梯度, 平滑分量梯度约束, 平滑分量Laplace约束, 由粗及精的多尺度控制策略.

双眼视觉是计算机视觉的重要方法, 双眼视觉的关键问题是对应点匹配问题, 基于特征的方法是解决对应点匹配问题的主要方法.

用图1表示立体视觉的成像模型. 通常与人双眼间的关系类似我们假设左右两个照相机之间的距离 B 较相机与景物间的距离小得多; 两个相机在竖直方向上是平齐的, 即右边相机可看作左边相机的水平平移, 且两个像平面处于同一平面内. 设景物表面上一点 A , 分别投射到左右两个像平面上的点 $A_L(x_L, y)$ 和点 $A_R(x_R, y)$ 上. 匹配主要由两个阶段来完成:

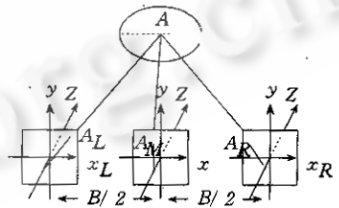


图1 成像模型及中点坐标系

算法 1. 传统方法

(1) 以一定的约束去寻找左(右)图中每一个特征点(边缘点) A_L 在右(左)图中的可能的匹配点.

(2) 选择出唯一的匹配点(此过程还要尽可能去除错误的匹配点).

* 本文1994-09-05收到, 1994-12-15定稿

作者钟声, 1966年生, 博士后, 主要研究领域为图象处理, 模式识别, 计算机视觉. 石青云, 女, 1936年生, 教授, 中国科学院院士, 主要研究领域为图象处理, 模式识别, 计算机视觉. 程民德, 1917年生, 教授, 中国科学院院士, 主要研究领域为应用数学, 模式识别, 计算机视觉.

本文通讯联系人: 石青云, 北京100871, 北京大学信息科学中心

第二步选择唯一匹配点,可能由于第一步选出过多的可能匹配点而需要过多的计算量和导致较多误匹配,算法的鲁棒性也就差了.因而更有效选择可能的匹配点是很有意义的.

许多工作致力于发展第二步选择唯一匹配点的算法,如松弛方法^[1]及建立符合物理或心理学上的依据的约束,如视差光滑性约束、视差梯度极限约束^[2],利用彩色信息^[3]等.还有的工作引入由粗及精的多尺度控制策略^[4,5].

建立更具鲁棒性的算法的关键问题之一是寻求适当的表示方法,使得在这种表示下双眼视觉信息融合所依据的特性得到很好的表示,噪声能得到较好的处理.此外,更具重要的是要能据此表示建立使双眼信息融合的物理上和心理上的约束条件,使得匹配问题得较好的解决.本文利用立体图对的二进小波变换表示,建立由粗及精的多尺度控制策略,以使不同尺度下的信息融合,容易引起误匹配的可能匹配点进一步减少,使得匹配算法的性能得以提高.本文第 1 节介绍二进小波变换及边缘提取,对噪声的处理及采用二进小波变换的理由也在本节中讨论;第 2 节给出了平滑分量梯度约束条件和平滑分量 Laplace 约束条件,第 3 节给出了匹配算法;第 4 节给出实验结果及与以往方法的比较;最后是小结和讨论.

1 二进小波变换,边缘检测与噪声处理

简单地讲,小波是一个积分(平均)为 0 的函数.设 $\psi(x)$ 是一小波函数,

定义 $\psi_{2^j}(x) = \frac{1}{2^j} \psi(\frac{x}{2^j})$, 则函数 $f(x)$ 在尺度 2^j 和位置 x 的小波变换定义为

$$W_{2^j}f(x) = f * \psi_{2^j}(x) = \int_{-\infty}^{+\infty} f(u) \psi_{2^j}(x-u) du \tag{1.1}$$

函数序列
$$Wf = \{W_{2^j}f(x)\}_{j \in Z} \tag{1.2}$$

就称为 $f(x)$ 的二进小波变换.

对于数字图像,其二维离散二进小波变换可以通过滤波器卷积运算实现.在下面的算法中,各符号的意义见文献^[6].

算法 2. 快速离散二进小波变换

给出一图象 S_1f :

$j=0$;

while ($j < J$) {

$$W_{2^{j+1}}^1 f = S_{2^j} * (G_j, D) / \lambda_j; \quad W_{2^{j+1}}^2 f = S_{2^j} * (D, G_j) / \lambda_j;$$

$$S_{2^{j+1}} f = S_{2^j} * (H_j, H_j); \quad j = j + 1;$$

}end of while

由上算法将图象 S_1f 分解到某一层级 J , 得到了 S_1f 的一个多尺度表示: $(\{W_{2^j}^1 f, W_{2^j}^2 f\}_{1 \leq j \leq J}, S_{2^j} f)$. 其中 $W_{2^j}^1 f$ 和 $W_{2^j}^2 f$ 分别是原图象的水平 and 垂直方向的高频(边缘)信息, 而 $S_{2^j} f$ 是原图的一个平滑的象. 实际上 $(\{W_{2^j}^1 f, W_{2^j}^2 f\}_{1 \leq j \leq J}, S_{2^j} f)$ 完备地描述了原信号 S_1f , 即由这些变换系数可以精确地重构原信号.

定义模
$$M_{2^j}f(x, y) = \sqrt{|W_{2^j}^1 f(x, y)|^2 + |W_{2^j}^2 f(x, y)|^2} \tag{1.3}$$

和角
$$A_{2^j}(x, y) = \text{复数}(W_{2^j}^1 f + iW_{2^j}^2 f) \text{的幅角}; \tag{1.4}$$

再检测沿 A_{2^j} 给出的方向上模 $M_{2^j}f(x, y)$ 的局部极大值点, $j=1, \dots, J$; 适当地选择小波, 这

样得到局部极大值点就给出了图象边缘的一个多尺度描述,该算法等价于 Canny 多尺度边缘检测算法^[7,8],而且 $A_{2^j}f(x,y)$ 描述了边缘点 (x,y) 梯度方向.

另外,(高斯白)噪声的二进小波变换还满足

$$E(|M_{2^j n}(x,y)|^2) = \frac{\sigma^2(\|\psi^1\|^2 + \|\psi^2\|^2)}{2^j}, \quad (1.5)$$

其中 E 表示数学期望值, σ^2 是 $n(x,y)$ 的方差, ψ^1, ψ^2 是小波函数^[9].

(1.5)式表明,当尺度增加一倍时, $|M_{2^j n}|^2$ 平均说来减小一半;所以在被噪声控制的点上,多尺度下边缘的描述可以利用(1.5)式来消除部分噪声的影响.

然后将边缘点连接成链.一个已链接边缘点的邻点中,将方向最接近、模值也大致相同的边缘点链接到边缘链中,去除链长小于某一阈值的链,因为它们通常是由噪声或一些微小的细节(容易引起误匹配)生成的.对立体图对进行上述一系列处理,得到了匹配算法的基本特征—左右图的边缘表示.图 2 和 3 中的(b)、(c)、(d)显示了原图对的前几个尺度 2,4,8 下的边缘.

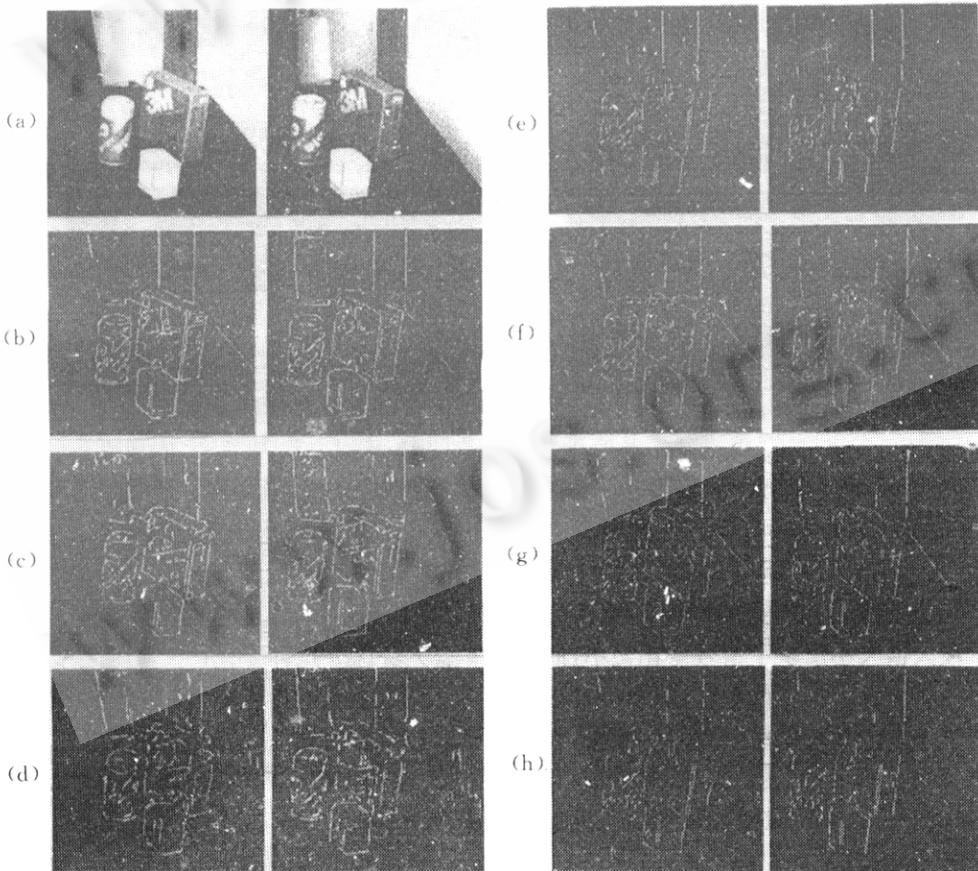


图 2

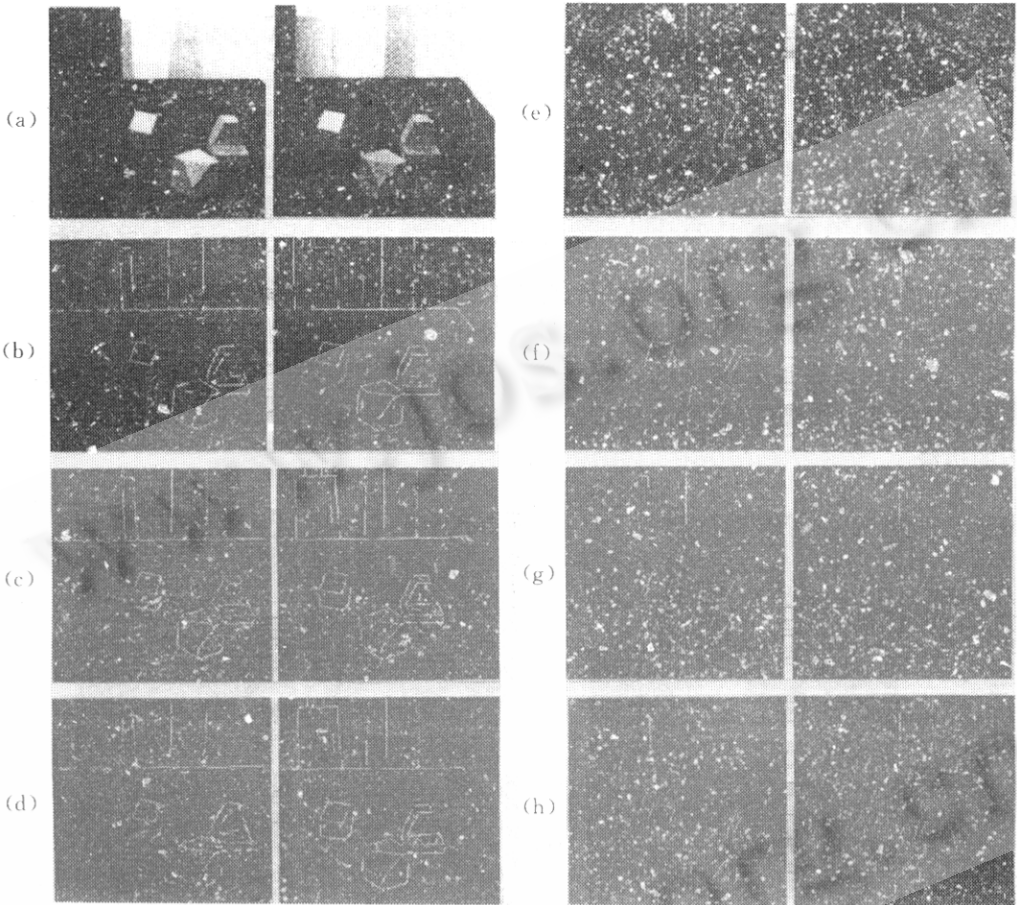


图 3

由多尺度极值链边缘及 $S_{2^j}f$ 可以重构原图的一个视觉效果相当好的近似图象^[8]. 这为我们基于二进小波变换的立体视觉方法提供了充分的依据, 因为从人类视觉理解的角度来看, 这表明多尺度边缘及 $S_{2^j}f$ 包含了足够多的理解图象的信息.

采用极值点检测较之过零点检测^[10]有几个好处:

- (1) 极大值大小反映了信号变化大小, 而用过零点无法区分这种变化;
- (2) 过零点仅给出了位置信息, 而极值点还可同时给出拐点的导数值, 特别是极大值点的值随尺度 s 的演化可以用来刻划奇异点(如边缘、噪声点)的性质.

一般情况下, (1.2)式提供的二进小波变换表示是冗余的. 对小波 ψ 加上一定条件, $\{\psi_{2^j}(2^j n - x)\}_{(n,j) \in \mathbb{Z}^2}$ 可以构成 $L^2(\mathbb{R})$ 的正交小波基^[11,12]. 此时,

$$\{W_{2^j}f(2^j n) = \langle f(x), \psi_{2^j}(2^j n - x) \rangle\}_{(n,j) \in \mathbb{Z}^2}$$

完全刻划了 $f(x)$, 即 $f(x)$ 完全由对 $W_{2^j}f(x)$ 进行 2^j 间隔的抽样决定. 对离散信号而言, 其正交小波变换是按 S. Mallat 的金字塔型分解^[11]进行的, 即每一层更进一步的分解都要进行二取一抽样. 十分值得注意的是, 这种抽样会给立体视觉问题造成不良的影响. 这是因为

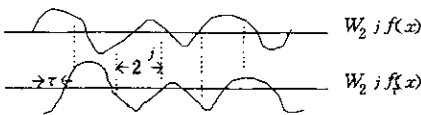


图4 平移与抽样的关系

当信号平移后重新计算的小波变换表示 $\{W_{2^j}(2^i n)\}_{(n,j) \in Z^2}$ 一般与对原信号计算的不一样 (而且不是只差一个平移的关系). 如图 4 所示, 下边信号为上边信号平移 τ 后得到的, 对连续的小波变换而言, 平移性是保持的, 即对平移后信号

的小波变换等于小波变换后的平移, 意即如果设 $f_\tau(x) = f(x - \tau)$ 为平移信号, 则

$$W_{2^j}(f_\tau(x)) = (W_{2^j}f)_\tau(x). \tag{1.6}$$

但是, 它们在 $2^i n$ 处的抽样却很不一样, 且不是平移关系, 如图 4 所示. 除非平移量 τ 恰好为 2^i 的倍数(此时, 抽样是一平移关系), 有

$$\{W_{2^j}f(2^i n)\}_{(n,j) \in Z^2} \neq \{W_{2^j}f_\tau(2^i n)\}_{(n,j) \in Z^2} \tag{1.7}$$

而在立体视觉中左右两幅图间的位置差异是显然的, 所以, 用正交小波变换来表征立体图对的信息是不合适的.

另一方面, 采用(1.2)式定义的二进小波变换, 边缘检测是通过变换系数的极大值点检测来完成的. 当信号平移 τ 时, $W_{2^j}f(x)$ 也平移 τ ((1.6)式), 因而极值点也平移 τ , 所以这种自适应的抽样方式可以保证立体图对中相应于景物同一特征的性质还在二进小波变换系数中得到相近的反映, 这正是匹配所要求和所依据的. 因而采用二进小波变换来表征立体视觉的图像信息是适当的.

2 平滑分量梯度约束和 Laplace 约束

多尺度极值链边缘表示还提供了许多用于指导双眼信息融合的条件.

首先, 传统的一些约束条件完全可由二进小波变换的系数来表达^[6]. 另外, 也是更重要的, 立体图对的多尺度极值链边缘和平滑分量包含了十分丰富的理解景物的信息, 应该还有更重要的更有效的帮助理解景物的约束条件存在, 推导这样的约束条件正是我们引入小波分析的真正原因. 在文献[6, 13]中, 我们推导出了下面几个约束条件.

依据双眼视觉成像模型, 如果左边一点 $A_L(x_L, y)$ 与右边一点 $A_R(x_R, y)$ 匹配, 那么两点的灰度值应该大致相等, 令 I_L, I_R 为左、右两图象的灰度函数, 则

$$I_L(x_L, y) = I_R(x_R, y). \tag{2.1}$$

定义一个新的坐标系: 如果匹配的 A_L 和 A_R 对应着景物表面的 A 点, 则 A 在新坐标系下的像平面(见图 1)上的投影为 $A_M(x, y)$, 其中

$$(x, y) = \left(\frac{x_L + x_R}{2}, y\right). \tag{2.2}$$

称这样定义的坐标系为中点坐标系. 则在中点坐标系下 $A_M(x, y)$ 点的视差为

$$h(x, y) = x_R - x_L; \tag{2.3}$$

中点坐标系下 A_L, A_R 的坐标可分别表示为 $(x - \frac{1}{2}h(x, y) - \frac{B}{2}, y)$ 和 $(x + \frac{1}{2}h(x, y) + \frac{B}{2}, y)$.

于是由(2.1)式得到

$$I_L(x - \frac{1}{2}h(x,y) - \frac{B}{2}, y) = I_R(x + \frac{1}{2}h(x,y) + \frac{B}{2}, y), \quad (2.4)$$

将左、右二图的二进小波变换的平滑分量 $S_{2^j}^L$ 和 $S_{2^j}^R$ 代入(2.4)式并在两边分别对 x 方向求一阶偏导数,再取绝对值等,可得

$$\begin{aligned} & |(S_{2^j}^L)_x(x - \frac{1}{2}h(x,y) - \frac{B}{2}, y) - (S_{2^j}^R)_x(x + \frac{1}{2}h(x,y) + \frac{B}{2}, y)| \\ & \leq \frac{\alpha}{2} |(S_{2^j}^L)_x(x - \frac{1}{2}h(x,y) - \frac{B}{2}, y) + (S_{2^j}^R)_x(x + \frac{1}{2}h(x,y) + \frac{B}{2}, y)|, \end{aligned} \quad (2.5)$$

$$\begin{aligned} & |(S_{2^j}^L)_y(x - \frac{1}{2}h(x,y) - \frac{B}{2}, y) - (S_{2^j}^R)_y(x + \frac{1}{2}h(x,y) + \frac{B}{2}, y)| \\ & \leq \frac{\alpha}{2} |(S_{2^j}^L)_y(x - \frac{1}{2}h(x,y) - \frac{B}{2}, y) + (S_{2^j}^R)_y(x + \frac{1}{2}h(x,y) + \frac{B}{2}, y)|, \end{aligned} \quad (2.6)$$

上式中 $(S_{2^j}^L)_x$ 等表示 $S_{2^j}^L$ 关于 x 的一阶偏导数等, α 是视差梯度分量绝对值 $\max\{|h_x|, |h_y|\}$ 的一个上界.

(2.5)与(2.6)式合称为平滑分量梯度约束^[13,14].

我们知道,图像的 Laplacian 变换(或者理解为二阶偏导数)包含了大量的信息,特别是与图像上灰度的变化程度有关,也即与边缘等视觉信息有关.因此,我们想进一步寻求立体图对的平滑分量的二阶偏导数满足的约束条件.

对于(2.4)式代入平滑分量后两边分别求 x, y 方向的二阶偏导数并经整理和取上界等推导,可以得到:

$$|(S_{2^j}^L)_{xx} - (S_{2^j}^R)_{xx}| \leq \alpha |(S_{2^j}^L)_{xx} + (S_{2^j}^R)_{xx}| + \frac{\beta}{2} |(S_{2^j}^L)_x + (S_{2^j}^R)_x|, \quad (2.7)$$

$$\begin{aligned} & |(S_{2^j}^L)_{yy} - (S_{2^j}^R)_{yy}| \\ & \leq \alpha |(S_{2^j}^L)_{yx} + (S_{2^j}^R)_{yx}| + \frac{\beta}{2} |(S_{2^j}^L)_x + (S_{2^j}^R)_x| + \frac{\alpha^2}{4} |(S_{2^j}^L)_{xx} - (S_{2^j}^R)_{xx}|, \end{aligned} \quad (2.8)$$

β 是 $|h_{xx}|$ 和 $|h_{yy}|$ 的一个上界. (2.7), (2.8) 式中, L, R 分量的坐标分别为 $(x - \frac{1}{2}h(x,y) - \frac{B}{2}, y)$ 和 $(x + \frac{1}{2}h(x,y) + \frac{B}{2}, y)$. 注意,在上面的推导中,我们假定视差和平滑分量都是足够光滑的.

我们将(2.7)、(2.8)合称为平滑分量 Laplacian 约束^[6,14].

特别值得一提的是参数 α, β 的取值. 按定义, α 是视差梯度 (h_x, h_y) 分量绝对值的一个上界, Burt & Julesz^[2] 的心理物理学实验证明了, 视差梯度是视场中景物能在双眼中融合的决定性因素. 当视差梯度的模值超过某一阈值如 1.0 时, 双眼的信息融合是不可能的. 另一方面在与人眼视觉系统相似的立体视觉系统中, 视差梯度 1.0 的限制包含了 95% 以上的情况. 所以, 无论在心理物理学上还是从统计学观点来看都是一个适当的选取. 另外, 从约束(2.5)–(2.8)的约束力来看, α, β 取值过大, 它们的约束力很弱, 使得很多不必要的可能匹配点入选, 既影响匹配精度, 又影响匹配速度; 反之, α, β 取值过小, 则它们的约束力太强, 有

可能使正确的匹配点被排除,也得不到正确的匹配.我们的实验也证明了 $\alpha = 1.0$ 、 $\beta \in [1.0, 2.0]$ 时它们的约束效力很高^[6,13-15].

约束(2.5)–(2.8)去除了大量可能引起误匹配的可能对应点而保留了正确的对应点,既提高了匹配的精度,又极大地提高了算法速度(平均一个数量级左右).

约束(2.5)–(2.8)在平滑分量上给出而不在原灰度函数上给出主要有两个原因,一是原图噪声较大,直接在原图上给出 Laplacian 约束和梯度约束往往得不到正确匹配;二是因为二进小波变换后,很多信息已经含在边缘表示中,再利用原图就有些重复了,容易使某些不利于匹配的因素被夸大,效果反而不好.实验也证实了这点.

平滑分量梯度约束和 Laplacian 约束的作用在于当左(右)图一边缘点在右(左)图的搜索范围内有若干个方向大致相同,模值相差不太大,但不满足平滑分量梯度约束和平滑分量 Laplacian 约束的边缘点时能将它们排除,从而减少了可能引起错误的因素,又提高了速度.

3 基于小波分析的多尺度立体视觉方法

我们已经建立了用于立体视觉的有效的表示——二进小波变换的多尺度边缘及用于匹配的有效的新的约束条件——平滑分量梯度约束和 Laplacian 约束.本节中,我们将综合前面的结果,建立一个由粗及精的多尺度立体视觉计算模型(算法).

这种由粗及精的多尺度的思想首先得益于大量生理和心理学的结论^[4].正如以前我们所提及的,在人类视觉系统处理视觉信息的某一阶段,视觉信息是在不同的频道中被独立地处理的;Marr 指出表面的灰度函数形成的空间结构通常是由不同的过程生成的,而且每个过程都是由一个不同的尺度下的操作来完成的.因此,即使仅从计算的角度出发,图像的结构也应该是在频率空间中来局部地确定.因而我们认为一个由粗及精的多尺度控制策略应被纳入立体视觉算法中.

算法 3. 基于小波分析的立体视觉方法

第一:利用二进小波变换提取立体图对多尺度极值链边缘(也包括平滑分量和幅角 $A_{2^j}^L, A_{2^j}^R, j=1, \dots, J$)(见第 1 节);

第二:对从最大的尺度 2^J 开始,到最小尺度 2^1 的每一尺度,对左(右)边图像上的每一个边缘点,确定其在右(左)边图像上的可能匹配点集,其中的每个可能的对应点满足:

- (1) 位于某一由控制策略决定的搜索范围中;(见由粗及精控制策略.四)
- (2) 有大致相同的方向(对 A_{2^j} 的约束)和相差不大的模值(对 M_{2^j} 的约束);
- (3) 满足平滑分量梯度约束(2.5)和(2.6);和
- (4) 满足平滑分量 Laplacian 约束(2.7)和(2.8).

第三:决定唯一匹配点,并消除误匹配.(利用 PMF 算法^[1]).

(1) 对每一对可能的匹配点 (i_L, i_R) , 计算来自其“支持邻域”的支持强度.(左(右)图上一边缘点的支持邻域定义为它的一个一定大小的邻域.)

如果该邻域中另外的边缘点(如 j_L)匹配后(其对应点为 j_R)使视差梯度小于 1.0, 即

$$\left| \frac{h\left(\frac{x_{i_R} + x_{i_L}}{2}, y_{i_L}\right) - h\left(\frac{x_{j_R} + x_{j_L}}{2}, y_{j_L}\right)}{\sqrt{(x_{i_L} - x_{j_L})^2 + (y_{i_L} - y_{j_L})^2}} \right| < 1.0, \quad (3.1)$$

也即

$$\left| \frac{(x_{i_R} - x_{i_L}) - (x_{j_R} - x_{j_L})}{\sqrt{(x_{i_L} - x_{j_L})^2 + (y_{i_L} - y_{j_L})^2}} \right| < 1.0, \tag{3.2}$$

那么邻点 j_L 对正在考察的匹配 (i_L, i_R) 的支持强度与 j_L 到 i_L 的距离成反比; 否则(即(3.1)或(3.2)不成立), 支持强度为 0; 匹配 (i_L, i_R) 从 i_L 的支持邻域中获得的支持强度为 i_L 的所有支持邻域中所有的边缘点对 (i_L, i_R) 的支持强度之和;

(2) 如果某一匹配 (i_L, i_R) 使得来自 i_L 支持邻域和来自 i_R 支持邻域的支持强度同时达到各自的可能匹配点集中的最大值, 则 (i_L, i_R) 被认为是被唯一匹配了;

(3) 对左、右图的所有边缘点进行了上述第三(1)、(2)步的处理后, 将所有可能匹配中凡是含有某一个已唯一匹配的点的可能匹配删除; 再进行第三(1)、(2)步的处理, 直到某一迭代次数;

第四: 由粗及精多尺度控制策略.

在最大尺度 2^J 下搜索范围为某一较宽的区间: $(x_L - w, x_L + w)$, 其中 w 约为最大估计视差模的 1.5 倍; 在较精细的尺度 2^j 下 ($j < J$), 对左(右)边图像上的每一个边缘点 (x_i, y_i) , 利用上一更大尺度 2^{j+1} 下得到的视差值, 在 (x_i, y_i) 的某个邻域中计算平均视差 d_i , 然后利用 $(x_i + d_i, y_i)$ 作为 2^j 尺度下的搜索中心, 搜索半径取与尺度 2^j 成正比的值且远比最大尺度 2^j 下的搜索半径 w 小. 在此过程中, 需注意处理异常情况: 当 2^{j+1} 尺度下的 (x_i, y_i) 附近的视差有突变(不连续)时, 需要对 (x_i, y_i) 的邻域作分割, 以便划出一个 (x_i, y_i) 的新邻域, 该邻域内有视差但没有突变, 再计算搜索中心(这样更准确); 如果 (x_i, y_i) 的一定范围邻域内在 2^{j+1} 尺度下没有视差值(即无匹配点), 则仍按最大尺度 2^j 下的搜索范围确定搜索空间.

在上面的算法中, 第三(1)中的支持邻域的引入实际上是引入了视差平滑约束, 邻域越大, 视差越平滑; (3.1)或(3.2)式称为视差梯度极限约束, 除了我们在第 2 节中提到的心理物理学和统计学上的意义外, 1.0 的限制还蕴含着某些几何意义上的约束, 如视差连续性约束, 有序约束(ordering constraint), 意即在一幅图像的某一水平上由左至右出现的特征, 在另一幅图像上也必须由左自右, 不能倒置. 整个确定唯一匹配点的方法(第三)是一种“Winner-take-all”的松弛过程, 此过程当第二步选出的候选匹配点较多时, 计算量很大, 且易产生误匹配. 引入第二中的各种约束条件及第四的由粗及精多尺度控制策略, 能很好地克服上述困难.

4 实验

我们在 DEC5000 工作站上用 C 语言编程做了许多实验, 证实了这个基于小波变换的多尺度双眼视觉匹配算法在速度和精度上都远优于传统的方法, 精度上较文献[6,13]的算法也有较大提高. 我们采用了二次样条小波函数^[8], 支持邻域的半径为 6.

图 2、3 的(a)图为立体图对, (b)~(d)为尺度 2, 4, 8 下的边缘, (e)、(f)、(g)、(h)的左图分别为由本文方法、文献[6]方法、传统方法(算法 1)和加上由粗及精多尺度控制策略的传统方法得到的视差图(disparity map), 右图是由相应的视差图恢复的右图像边缘位置图; 视差图是将实际视差按比例变为可显示的图像, 亮度越大表示视差越大, 反之则越小. 图 3 的

噪声很严重.

表 1、表 2 给出了一些统计结果. 其中,对本节算法列出的是最后一次迭代的结果,传统算法与文献[6]的算法未采用由粗及精多尺度控制策略,且对水平边缘点也作了匹配,所以最后被唯一匹配的点数多些.

表 1 对图 2 的实验统计数据

算法	本文算法	文献[6]的算法	多尺度传统算法	传统算法 1
有候选匹配点的左边缘点数	1344	1685	1447	1924
可能的匹配点总数	4395	7045	7036	12941
最后被唯一匹配的点数	1177	1511	1068	1389
误匹配情况	几乎没有	少数	一些	较多
匹配过程的系统时间(秒)	18	11	51	38

表 2 对图 3 的实验统计数据

算法	本文算法	文献[6]的算法	多尺度传统算法	传统算法 1
有候选匹配点的左边缘点数	456	675	495	1056
可能的匹配点总数	1411	2284	1718	11411
最后被唯一匹配的点数	390	544	405	595
误匹配情况	几乎没有	少数	一些	较多
匹配过程的系统时间(秒)	12	5	27	71

表中对误匹配的检查是靠人机交互检验的,我们没给出具体数据,但从视差图已经可以看出表中的结论,图中一个物体轮廓线上亮度突变的地方是由误匹配导致,可以发现传统方法视差图中的亮度突变的地方多得多. 表中时间未包括边缘提取所需的系统时间(约 20 秒).

本文的方法与传统方法(算法 1)相比可以看到如下几个结论:(1)可能的匹配情况总数大大减少,而最后唯一匹配的点的数目却相当;(2)精度提高极大(以误匹配数目少来度量);(3)匹配所需的时间大大减少;(4)算法对噪声环境更稳定(特别注意对图 3 的匹配结果).

所以平滑分量 Laplacian 约束和平滑分量梯度约束一起去除了大量可能引起误匹配的点对应点而保留了正确的对应点,既提高了初始匹配的精度,又极大地提高了算法速度;由粗及精多尺度控制策略与平滑分量 Laplacian 约束和平滑分量梯度约束一起进一步融合了正确的对应点而排除了错误匹配.

本文的方法与采用由粗及精多尺度控制策略的传统方法相比也要好得多.

与文献[6,13]中的算法相比可以看到如下几个结论:(1)匹配精度提高了(以误匹配数目少来度量),特别注意完整的轮廓线得到了较好的恢复;(2)算法对噪声环境更稳定. 所以由粗及精多尺度控制策略可以提高匹配的精度,使算法鲁棒性增强.

5 讨论和小结

本文我们利用二进小波变换建立了立体视觉的一个高效的计算模型. 首先, 利用小波分析完成原始图像(立体图对)到基本要素图(多尺度边缘及平滑分量)的转换. 其次, 我们建立了多个有效的对候选匹配点的约束条件——平滑分量梯度约束和 Laplacian 约束, 它们的建立, 既符合人类视觉系统的心理学机制, 又符合统计学上的规律, 因而发挥了十分有效的作用. 另外, 我们利用人类立体视觉多频道处理的知识, 建立了一个由粗及精的多尺度立体视觉匹配算法, 获得相当好的结果.

这些好效果的取得, 使得我们更加坚信: 小波分析在某些方面符合人类视觉处理的机制, 特别是在其表征视觉信息的能力上, 显现出了较以往方法优越得多的特点. 例如, 我们建立的几个约束及传统的一些约束, 都在小波变换域中得到自然的表达.

在小波分析所表征的视觉信息的利用上, 还有许多潜力可挖, 例如 Lipschitz 系数^[8]可用以刻画边缘的类型, 如区分某边缘是由遮挡或是阴影造成, 所以可利用 Lipschitz 系数建立匹配的约束条件^[15]. 关于这方面的研究, 我们认为一个重要的方向就是从充分了解人类双眼视觉机制入手, 以便更主动、更有意识地利用小波分析表征视觉信息. 另外, 从立体视觉计算的角度来讲, 还可做许多进一步的研究, 例如用于匹配的特征可以不是边缘点, 而是边缘段(即轮廓线段、曲线段)等. 再者, 还可考虑更一般的双眼视觉模型. 我们今后将进行这些方面的研究.

致谢 北京大学迟惠生教授、仇桂生教授与彭立中教授给予作者大量指教; 北方交通大学信息所的袁保宗教授为我们提供了立体图对; 另外视觉听觉信息处理国家重点实验室的老师也给了许多帮助, 在此一并致谢.

参考文献

- 1 Pollard S B, Mayhew J E W, Frisby J P. PMF: a stereo correspondence algorithm using a disparity gradient limit. *Perception* 14, 1985, 449—470.
- 2 Burt P, Julesz B. A disparity gradient limit for binocular fusion. *Science* 208, 1980, 615—617.
- 3 Jordan J R, Bovik A C. Using chromatic information in edge-based stereo correspondence. *CVGIP: Image Understanding*, 1991, 54(1), 98—118.
- 4 Marr D. *Vision*. San Francisco: Freeman 1982.
- 5 Grimson W. Computational experiments with a feature based stereo algorithm. *IEEE Trans. on PAMI*, 1985, 7(1):17—34.
- 6 Zhong Sheng, Shi Qingyun, Cheng Min—Teh. Improved stereo correspondence algorithm based on wavelet transform. *Proceedings of Asian Conference on Computer Vision, Japan, 1993*. 268—271. (中文稿见《软件学报》, 1994, 5(8):1—8.)
- 7 Canny J. A computational approach to edge detection. *IEEE Trans. on PAMI*, 1986, 8:679—698.
- 8 Mallat S, Zhong Sifen. Characterization of signals from multiscale edges. *IEEE Trans. on PAMI*, 1992, 14(7).
- 9 Mallat S, Hwang W L. Singularity detection and processing with wavelets. *IEEE Trans. on IT*, 1992, 38(2).
- 10 Marr D, Hildreth E. Theory of edge detection. *Proc. Roy. Soc. London B*, 1980, 207:187—216.
- 11 Mallat S. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Trans. on PAMI*,

- 1989, 11(7):674—693.
- 12 Daubechies I. Orthonormal bases of compactly supported wavelets. Commun. on Pure & Applied Mathematics, 1988, XLI:909—996.
- 13 Zhong Sheng, Shi Qingyun, Cheng Min—Teh. Stereo vision based on wavelet transform. Proceeding of the IEEE Signal Processing Society 8th Workshop on Image & Multidimensional Signal Processing, Cannes, France, Sept. 1993. (中文稿见《模式识别与人工智能》, 1994, 7(1):27—33.)
- 14 钟声. 小波变换及其在图像处理 and 计算机视觉的应用[博士学位论文]. 北京大学, 1994.
- 15 Zhong Sheng, Shi Qingyun, Cheng Min—Teh. Using lipschitz regularity in stereo vision. to be published.

MULTISCALE STEREO VISION BASED ON WAVELET TRANSFORM

Zhong Sheng Shi Qingyun Cheng Minde

(Department of Mathematics, Beijing University, Beijing 100871)

(The Center of Information Science, Beijing University, Beijing 100871)

(The National Laboratory on Machine Perception, Beijing University, Beijing 100871)

Abstract This paper presents an efficient multiscale binocular correspondence algorithm based on dyadic wavelet transform. The stereo images are first decomposed on the dyadic wavelet bases. By analyzing the transform coefficients, the authors get an efficient method to remove noises and extract edges. Some basic constraints for matching, e. g. the one which requires that the edge directions be nearly the same, are naturally expressed by the coefficients. More importantly, they can deduce four useful constraints for selecting the potential matches which are called the smooth component gradient constraints and the smooth component Laplacian constraints. They have strong disambiguation power and greatly remove those potential matches that would cause mismatches, and thus greatly improve the accuracy and speed for matching. A multiscale coarse to fine control strategy is included to increase robustness.

Key words Stereo vision, correspondence problem, dyadic wavelet transform, disparity gradient, smooth component gradient constraints, smooth component Laplacian constraints, coarse to fine control strategy.