

改进的基于小波变换的立体视觉匹配方法*

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双眼视觉匹配算法. 与传统的算法相比, 该算法具有更高的匹配精确度和更快的匹配速度(提高1个数量级左右). 与作者过去提出的方法相比, 本文又推导出两个有效的新的约束——平滑分量Laplace约束; 它们与平滑分量梯度约束一起, 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. 另外, 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, 提出的边缘将作为对应点匹配的基本特征. 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. 本方法是解决立体视觉匹配问题的一个十分有效的模型.

关键词 计算机视觉, 立体匹配, 对应点问题, 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.

计算机视觉的最重要的方法之一是立体视觉方法, 尤以双眼视觉为代表. 较之基于单幅图象的视觉方法(如由明暗度或线画图确定形态), 双眼视觉是从同一景物投射到两个象平面所得到的两幅图象来确定景物的形状及其距视点的距离, 它是可定位的——通过寻找两幅图象上的对应点. 因此, 双眼视觉方法具有重要的实际意义.

双眼视觉的关键问题是对应点问题, 即寻找两幅图象上对应于同一景物点的象点. 该问题的主要困难在于景物中各物体的空间位置关系复杂而造成相互遮拦、噪声的影响等等. 另外, 透明物体的出现也带来了极大困难. 由于上述原因及考虑到实际计算上的消耗, 对左(右)边图象上所有点寻找右(左)边图象上的对应点是不实际和困难的, 因而通常我们只能寻找一些有代表性的点的对应点, 而且这样的代表点应该足够稠密, 不然我们从稀疏的对应点匹配中得不到景物中各物体的较精确的形状和距离. 另外, 还要求这些代表点能提供稳定的匹配结果, 即对大多数景物都有较好的匹配结果及抗噪声能力较强.

物体的边缘正好是满足要求的代表点, 边缘还在人类视觉理解中起着重要的作用^[1].

* 本文1993-01-06收到

本项目部分受国家自然科学基金的支持. 作者钟声, 28岁, 博士生, 主要研究领域为图象处理, 模式识别, 计算机视觉. 石青云, 女, 58岁, 中国科学院院士, 教授, 主要研究领域为图象处理, 模式识别, 计算机视觉. 程民德, 77岁, 中国科学院院士, 教授, 主要研究领域为模式识别, 图象处理, 泛函分析.

本文通讯联系人: 钟声, 北京 100871, 北京大学信息科学中心

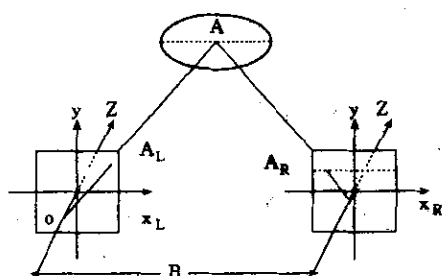


图1 成像模型

图1给出了景物中一点A向两个象平面(照相机)投射成像的几何原理.通常,与人的双眼间的关系类似,我们假定两个照相机的光轴是平行的,象平面处于同一平面内,两个相机中心之间距离B远远小于相机到景物的距离.加上外极线约束,即相机底线处于同一直线上,从而左右两幅图象在竖直方向上是平齐的,这样寻找 A_L 的对应点 A_R 就只需要到右图中与 A_L 具有相同纵坐标的直线上去搜索,只需求出视差 $x_R - x_L$ 即可.

在此模型下基于边缘的立体视觉匹配方法主要由两个阶段来完成.首先,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;然后再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).其中,第一步对选择的可能匹配点传统的约束有:(i)位于某一适当的搜索范围内;(ii)与 A_L 具有相同的对比符号;(iii)与 A_L 具有基本相同的方向.第二步选择唯一匹配点,可能由于第一步选出的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.因而如何更有效地选择可能的匹配点是很有意义的.作者在文献[7]中利用小波变换,推导出了可以大量减少容易引起误匹配的可能匹配点的约束条件——平滑分量梯度约束,更充分地利用了图象提供的信息,得到了满意的结果.本文中我们又发现了两个有效的约束条件——平滑分量Laplace约束,使得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进一步提高.

本文第1节介绍二进小波变换及边缘提取,对噪声的处理也在本节中讨论.第2节推导出了平滑分量梯度约束条件和平滑分量Laplace约束条件,并给出了匹配算法.实验结果及与以往方法的比较在第3节给出.最后是小结和讨论.

1 二进小波变换、噪声处理及边缘提取

小波变换近年来引起了数学家和工程学家极大重视.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

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

函数叙列 $Wf = \{W_{2^j}f(x)\}_{j \in \mathbb{Z}}$ 就称为 $f(x)$ 的二进小波变换.实际上, Wf 是函数 $f(x)$ 的一个多尺度刻划^[3].

下面我们给出一个快速二维小波变换的离散算法.我们用 $S * (A, B)$ 表示这样的运算:图象S的每一行与一维滤波器A作卷积,再将结果的每一列与一维滤波器B作卷积.下面的算法中, D 是Dirac滤波器; G, H 分别是一维的高通和低通滤波器,它们是由适当的小波函数推导出来的; G_j, H_j 表示分别在 G 和 H 的相邻系数间插入 $2^j - 1$ 个零而得到的离散的滤波器, λ_j 是正则化因子.

算法 1. 快速离散二进小波变换

```

    给出一图象  $S_1 f$ ;
     $j=0$ ;
    while( $j < J$ ) {
         $W_{2^j+1}^1 f = S_{2^j} * (G_j, D) / \lambda_j$ ;     $W_{2^j+1}^2 f = S_{2^j} * (D, G_j) / \lambda_j$ ;
         $S_{2^{j+1}} f = S_{2^j} * (H_j, H_j)$ ;     $j=j+1$ ;
    } end of while
  
```

由上算法将图象 $S_1 f$ 分解到某一层次 J , 得到了 $S_1 f$ 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 $S_1 f$, 即由这些变换系数可以精确地重构原信号.

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

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

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就给出了图象边缘的一个多尺度描述^[3], 由此得到的边缘算法等价于 Canny 多尺度边缘检测算法^[4], 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

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

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^[5].

(2) 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(2)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 3b—图 5b 显示了原图对的边缘(我们可以认为前几个尺度如 2, 4, 8 下, 边缘的位置相同^[3], 图中显示的是尺度 4 下的边缘).

2 匹配算法

立体视觉图对的二进小波变换不仅使我们得到了用于匹配的基本特征——边缘表示, 噪声也得到了恰当的处理, 而且使我们发现了一些更有用的约束条件.

首先, 还是让我们来看一看二进小波变换的多尺度极值链边缘及平滑分量 $S_{2^J} f$ 所包含的信息.

S. Mallat 等^[3]验证了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. 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.

那么, 我们从 $S_{2^J} f$ 能得到什么可以进一步去除容易引起误匹配的可能对应点的约束

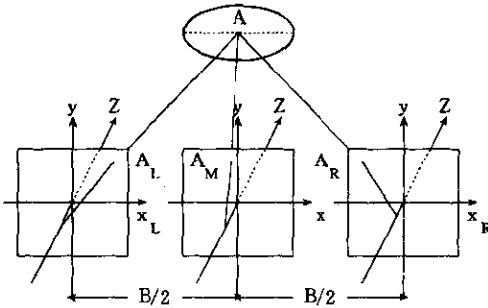


图2 中点坐标系

呢?这种约束同时还能匹配的后一阶段选取唯一匹配点节省大量计算。

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

$$I_L(x_L, y) = I_R(x_R, y) \quad (3)$$

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2)上的投影为 $A_M(x, y)$,其

中, $(x, y) = (\frac{x_L + x_R}{2}, y)$. 称这样定义的坐标系为中点坐标系。

则在中点坐标系下 $A_M(x, y)$ 点的视差定义为 $h(x, y) = x_R - x_L$; 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(3)式得到 $I_L(x - \frac{1}{2}h(x, y) - \frac{B}{2}, y) = I_R(x + \frac{1}{2}h(x, y) + \frac{B}{2}, y)$ (4)

将左、右两图的二进小波变换的平滑分量 S_2^L 和 S_2^R 代入(4)式也应成立;两边分别对 x , y 方向求一阶导,可以推出

平滑分量梯度约束:

$$\begin{aligned} & |(S_2^L)_x(x - \frac{1}{2}h(x, y) - \frac{B}{2}, y) - (S_2^R)_x(x + \frac{1}{2}h(x, y) + \frac{B}{2}, y)| \\ & \leq \frac{\alpha}{2} |(S_2^L)_x(x - \frac{1}{2}h(x, y) - \frac{B}{2}, y) + (S_2^R)_x(x + \frac{1}{2}h(x, y) + \frac{B}{2}, y)|, \end{aligned} \quad (5)$$

$$\begin{aligned} & |(S_2^L)_y(x - \frac{1}{2}h(x, y) - \frac{B}{2}, y) - (S_2^R)_y(x + \frac{1}{2}h(x, y) + \frac{B}{2}, y)| \\ & \leq \frac{\alpha}{2} |(S_2^L)_y(x - \frac{1}{2}h(x, y) - \frac{B}{2}, y) + (S_2^R)_y(x + \frac{1}{2}h(x, y) + \frac{B}{2}, y)|, \end{aligned} \quad (6)$$

和平滑分量 Laplace 约束:

$$|(S_2^L)_{xx} - (S_2^R)_{xx}| \leq \alpha |(S_2^L)_{xx} + (S_2^R)_{xx}| + \frac{\beta}{2} |(S_2^L)_x + (S_2^R)_x|, \quad (7)$$

$$|(S_2^L)_{yy} - (S_2^R)_{yy}| \leq \alpha |(S_2^L)_{yy} + (S_2^R)_{yy}| + \frac{\beta}{2} |(S_2^L)_y + (S_2^R)_y| + \frac{\alpha^2}{4} |(S_2^L)_{xx} - (S_2^R)_{xx}|, \quad (8)$$

其中 $(S_2^L)_x$, $(S_2^L)_{xx}$ 分别表示沿 x 方向的一阶导数和二阶导数, α 是视差梯度分量绝对值 $|h_x|$, $|h_y|$ 的一个上界, α 通常取值 1.0, 这是有生理物理学依据的^[6], 只有满足视差梯度 1.0 的约束, 双眼视觉信息才能融合; 也有统计学上的依据, 1.0 对大多数情形是适用的(可以证明 90% 以上的情形^[6]). 如果 α 值过大, 则(5)~(8)的约束力很弱, 得到的候选匹配点越多, 既影响匹配的精度又影响速度; 反之, α 过小, (5)~(8)约束力强, 得到的候选匹配点少, 可能去掉了正确的匹配点, 影响匹配精度. β 是 $|h_{xx}|$ 和 $|h_{yy}|$ 的一个上界, 1.0~2.0 是 β 实验上的合适的取值范围; 之所以引入(7)(8)约束, 是考虑到一个信号的 Laplace 变换包含了大量的视觉信息。

由此, 我们对可能匹配的对应点对加上(5)~(8)4 个约束条件, 得到下述的匹配算法。

算法 2. 双眼视觉匹配算法(基于二进小波变换):

一、对左(右)边图象的每一个边缘点,确定右(左)图的可能的匹配点集,其中每个可能的对应点满足:

(1)位于某一适当的搜索范围中;(2)有大致相同的方向(对 A_2' 的约束)和相差不大的模值(对 M_2' 的约束);(3)满足平滑分量梯度约束(5)和(6);和(4)满足平滑分量 *Laplace* 约束(7)和(8)。

二、再利用 *PMF* 算法来决定唯一的匹配,并消除误匹配。

(1)对每一可能的匹配计算其来自“支持邻域”的支持值. 左(右)图一边缘点的支持邻域为它的一个邻域. 如果该邻域中另外的边缘点匹配后使正在考虑的可能匹配满足视差梯度极限 1.0 的约束,则该邻点对正在考虑的可能匹配的支持与邻点到该点的距离成反比;否则支持为 0;(2)如果某一匹配使得左右两边的点同时获得最大的支持,则这一点对被认为是被唯一匹配了;(3)对左、右图的边缘点都进行上面 1、2 的处理后,将含有被唯一匹配的点的所有可能匹配去掉,再进行上述 1、2 的处理,直至某一迭代次数。

此算法中一(1)、(2)相当于传统算法(见第 1 节)的(i)(ii)(iii);二(1)中支持邻域的支持相当于作了视差平滑约束,邻域越大,视差越平滑。

约束(5)–(8)的作用在实验中得到了验证。

3 实验结果

我们在 DEC 5000 工作站(UNIX 操作系统)上用 C 语言实现了整个算法。

我们做了许多实验,证实了基于小波变换的双眼视觉匹配算法在速度和精度上都优于传统的方法。

我们采用了二次样条小波函数^[4]. 支持邻域的半径为 6。

立体图对是由一架 CCD 照相机平移一段距离后得到的. 图 3(a)、4(a)、5(a)是 $256 \times 256 \times 8\text{bits}$ 的原立体图对,3(b)、4(b)、5(b)是由二进小波变换提取的边缘表示(可以认为前几个尺度如 2, 4, 8 下,边缘的位置相同^[3]),噪声已经大大减少了. 原图对都含有噪声干扰,图 5(a)尤为严重。

为了便于比较,我们同时也给出传统匹配算法的结果,即不要约束(5)–(8)。

图 3(c)、4(c)、5(c)的左图为由本文的方法得的视差图(disparity map),右图为传统方法得到的视差图,视差图是将实际视差按比例变为可显示的图象,亮度越大表示视差越大,反之则越小。

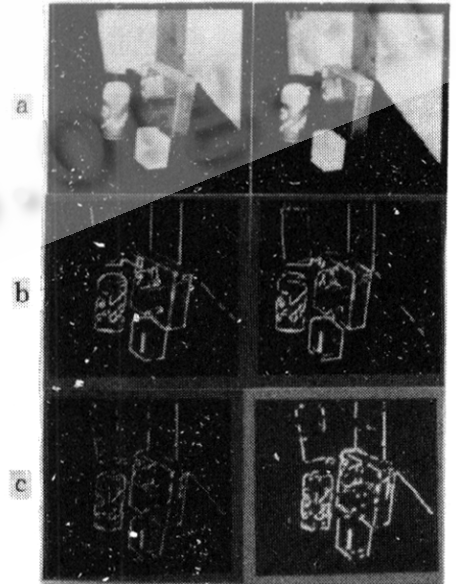


图 3

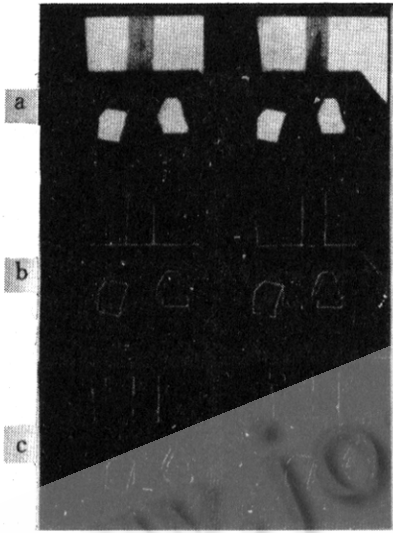


图 4

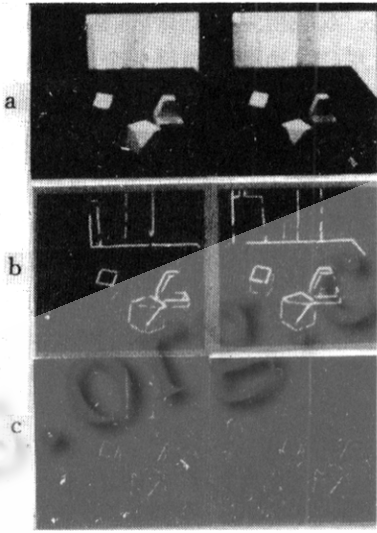


图 5 (a)原图;(b)边缘;

(c)视差图(左:本文方法;右:传统方法).

匹配是在尺度 4 下进行的,因为我们认为从主要视觉信息保持的角度来看,前几个尺度如 2,4,8 下,边缘的位置相同^[3];实际上,不同尺度下的边缘是有差别的,我们可以采用由粗及精的多尺度匹配方法,得到更好的结果^[8].表 1 给出了一些统计结果.

表 1 实验统计数据

	图 2		图 3		图 4	
	有(5-8)约束	无(5-8)约束	有(5-8)约束	无(5-8)约束	有(5-8)约束	无(5-8)约束
有候选匹配点的左边缘点数	1709	1893	667	829	698	954
可能的匹配点总数	6938	13326	2058	11165	2407	12047
最后被唯一匹配的点数	1593	1518	635	632	618	708
误匹配情况	很少	很多	很少	很多	很少	很多
匹配过程的系统时间(秒)	11	41	4	89	5	94

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.表中时间未包括边缘提取所需的系统时间(大约 20 秒).

本文的方法与传统方法相比有如下几个结论: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;说明约束(5),(6)排除了大量不必要的候选对应点.2. 精度提高了(以误匹配数目少来度量).3. 匹配所需的时间大大减少了,甚至减少 1 个数量级以上.4. 算法在噪声环境里是稳定的(性能良好).

所以约束(5)-(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。

与文献[7]的方法比较,加入(7)(8)约束后,匹配的精度和速度又有提高,不因(7)(8)的计算而减慢,这是因为(7)(8)进一步减少了可能引起误匹配的可能对应点,使得匹配的第二阶段节省时间并减少误匹配。

注意,我们的实际处理中,对水平边缘点也进行了匹配,对大多数这类边缘点来说,匹配的结果较之传统算法仍是较精确的。

4 小结与讨论

我们利用二进小波变换来求解双眼视觉的对应点匹配问题。二进小波变换不仅提供了图象的边缘信息作为匹配算法的基本特征,而且提供了平滑分量梯度约束和平滑分量 Laplace 约束四个强有力的约束条件,同时对噪声也做出了适当的有效的处理。

那么约束(5)–(8)为什么要在平滑分量上给出而不在原灰度函数上给出呢?主要有两个原因,一是原图噪声较大,直接原图上给出约束(5)–(8)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。

约束(5)–(8)的作用与传统方法比在于当左(右)图 1 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(5)–(8)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。

目前,本方法是基于边缘点的匹配,更高效的方法是进一步利用一些关于物体结构方面的知识和人类视觉系统的知识,如基于轮廓线的方法和多尺度的方法,这是我们正在开展的工作。

致谢 作者十分感谢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,仇桂生教授与彭立中教授也给予大量指教;另外许公望等同学及视觉听觉国家重点实验室的崔立依等老师也给了许多帮助,一并致谢。

参考文献

- 1 Marr D. Vision. San Francisco: Freeman, 1982.
- 2 Pollard S B, Mayhew J E W, Frisby J P. PMF: a stereo correspondence algorithm using a disparity gradient limit. Perception, 1985;93–99.
- 3 Mallat S, Zhong Sifen. Characterization of signals from multiscale edges. IEEE Trans. on PAMI, 1992,14(7).
- 4 Canny J. A computational approach to edge detection. IEEE Trans. on PAMI, 1986,8:679–698.
- 5 Mallat S, Huang W L. Singularity detection and processing with wavelets. IEEE Trans. on IT, 1992,38(2).
- 6 Bart P, Jules B. A disparity gradient limit for binocular fusion. Science, 1980,206:615–617.
- 7 Zhong Sheng, Shi Qingyun, Cheng Minde. Stereo vision based on wavelet transform. Proceedings of the IEEE Signal Processing Society 8th Workshop on Image and Multidimensional Signal Processing, Cannes, France, Sept. 1993.
- 8 Zhong Sheng, Shi Qingyun, Cheng Minde. Multiscale stereo vision based on wavelet transform. Submitted to 12th ICPR, Tel Aviv, Oct. 1994.

AN IMPROVED STEREO CORRESPONDENCE ALGORITHM BASED ON WAVELET TRANSFORM

Zhong Sheng, Shi Qingyun and Cheng Minde

(Department of Mathematics, Peking University, Beijing 100871)

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

(State Key Laboratory on Machine Perception, Peking University, Beijing 100871)

Abstract This paper presents an efficient binocular correspondence algorithm based on dyadic wavelet transform. Compared with the conventional algorithm, this method achieves higher accuracy and speed (improved by about one order of magnitude). 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 they call 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.

Key words Computer vision, stereo vision, correspondences, dyadic wavelet transform, disparing gradient, smooth components gradient constraints, smooth components Laplace constraints.