

基于 Quaternion-Gabor 小波运动估计的深度图获取技术

罗桂娥, 徐运彬*

(中南大学 信息科学与工程学院, 长沙 410083)

(* 通信作者电子邮箱 yunbinxu@sina.cn)

摘要:深度图是“二维视频+深度图”格式 3D 显示器的关键技术所在。在对四元数及 Gabor 滤波器进行研究的基础上提出了基于 Quaternion-Gabor 小波运动估计的深度图获取技术。通过计算普通视频运动图像的全局运动向量,估计出图像背景运动模型,获得图像运动场,最后分离出图像的前景与背景,从而得到图像的深度图。将普通的 Gabor 滤波器扩展到 Quaternion-Gabor 滤波器,不仅把图像变换到频率域获得额外的信息,而且可以对每个像素的 RGB 分别独立滤波。实验结果表明,使用 Quaternion-Gabor 小波运动估计方法得到的深度图变化很平滑,边缘更突出。

关键词:二维视频;深度图;四元数;小波;运动估计

中图分类号: TP391.413 **文献标志码:** A

Depth map acquisition technique based on Quaternion-Gabor wavelet motion estimation

LUO Gui-e, XU Yun-bin*

(School of Information Science and Engineering, Central South University, Changsha Hunan 410083, China)

Abstract: Depth map is the key technology of “2D video + depth map” for 3D display. On the basis of the research into quaternion and Gabor filter, the depth map acquisition technique based on Quaternion-Gabor wavelet motion estimation was proposed. Through calculating the global motion vector of image from ordinary video, the background motion model was estimated and the motion field was gotten. In the end, the foreground and the background of the image were isolated, and the depth map of the image was obtained. Through expanding ordinary Gabor filter to Quaternion-Gabor filter, it can not only get extra information through transforming the picture to frequency domain, but also can get independent filter to each pixel's RGB. The experimental results show that the changes of depth map obtained by Quaternion-Gabor wavelet motion estimation will be very smooth and edges will be more outstanding.

Key words: 2D video; depth map; quaternion; wavelet; motion estimation

0 引言

随着飞利浦公司“二维视频+深度图”格式 3D 显示器的问世,只要在原有二维视频的基础上引入对应的深度图,就能在该显示器中播放出 3D 效果^[1]。目前影视拍摄采用普通摄像机,没有深度信息的获取功能。如在采集端获取深度信息必然会导致产业链的更新换代,推广难度极大。对早期已制作好的影视作品,同样无法通过采集端获取深度信息^[2]。所以,将二维视频转换成三维视频是目前解决立体影视内容匮乏以及解决立体产业发展瓶颈的一个非常重要的手段^[3]。

目前单目深度图提取方法主要有线性透视方法、利用物体边缘获得深度信息的方法、基于离焦的方法、基于聚焦的方法、基于大气透视的方法和基于运动估计的方法等^[4]。众所周知,大家所观看的视频以运动场景为主,所以利用运动估计获得深度信息才是直接与物体的真实深度信息相对应的。

本文提出了利用 Quaternion-Gabor 小波滤波技术,对视频前后帧进行匹配,从而获得场景的相对深度信息的方法。该方法获取的深度信息更加精确,而且深度图让人看起来更加舒服。

1 四元数简介

早在 18 世纪,著名的数学家 Hamilton 在研究三倍数 (Triple Number) 乘法时发明了四元数 (Quaternion)^[5]。19 世纪初期,Albert Einstein 为了让光速对所有的惯性观察者来说都是常量,时间和空间必须统一,这样就发现了四元数的第一个应用实例^[6]。

在 Clifford 代数中,超复数是代数学家 Clifford 对那些比一般复数具有更高维的复数命名,它具有实数和复数的性能。最常见的超复数包括双四元数 (Biquaternions)、八元数 (Octonions)、四元数 (Quaternions)。Hamilton 发明的四元数是利用超复数定义的一个体系,它的乘法具有不可交换性^[7]。

四元数可以用超复数形式表示:

$$q = a + bi + cj + dk \quad (1)$$

其中: a, b, c, d 均为实数; i, j, k 是复数算子,遵循下列规定:

$$ijk = i^2 = j^2 = k^2 = -1 \quad (2)$$

$$jk = i, kj = -i, ki = j, ik = -j, ij = k, ji = -k \quad (3)$$

四元数由 1 个实部和 3 个虚部组成。虚部可以进行向量运

收稿日期:2011-07-18;修回日期:2011-08-16。

基金项目:国家自然科学基金重大专项(90820302);国家博士点基金资助项目(200805330005)。

作者简介:罗桂娥(1962-),女,湖南株洲人,教授,博士研究生,主要研究方向:图像信号处理、立体视频信号处理、智能仪器;徐运彬(1986-),男,湖南邵阳人,硕士研究生,主要研究方向:数字图像处理、软件设计。

算,记为:

$$\mathbf{V}(q) = bi + cj + dk \quad (4)$$

设用 $S(q)$ 表示四元数的实部,即标量部分,则整个四元数可以用标量和向量两部分表示^[8],记为:

$$q = S(q) + \mathbf{V}(q) \quad (5)$$

如果一个四元数的实部或标量部分为0,则称为纯四元数。

四元数的 Cayley-Dickson 表示形式^[8]为:

$$q = A + B\mathbf{j} \quad (6)$$

其中: $A = a + bi, B = c + di$,使得:

$$q = (a + bi) + (c + di)\mathbf{j} \quad (7)$$

采用正交算子 i 和 j 的规则,将式(7)乘出来,得:

$$q = a + bi + cj + dk \quad (8)$$

2 Quaternion-Gabor 滤波器

2.1 Gabor 滤波器

Gabor 滤波器可以看成经典傅里叶变换的截断形式。利用傅里叶变换进行图像处理时可以把图像变换到频率域,这样可以获得更多额外的信息。在传统的傅里叶变换中,只能对图像进行全局处理,而 Gabor 滤波器在处理局部特征时有很好的效果。

二维 Gabor 滤波器基函数的一般形式给出了定义:

$$h(x, y) = g(x', y') \exp[-2\pi i(U + V)] \quad (9)$$

其中:

$$\begin{cases} x' = x \cos \theta + y \sin \theta \\ y' = -x \sin \theta + y \cos \theta \end{cases} \quad (10)$$

x', y' 是坐标 x, y 角度 θ 旋转后的坐标表示。其中:

$$g(x, y) = \frac{1}{2\pi\delta} \exp\left(-\frac{x^2 + y^2}{2\delta^2}\right) \quad (11)$$

由式(11)可知, $g(x, y)$ 是一个二维低通高斯滤波器。而 Gabor 滤波器 $h(x, y)$ 是一个被正弦函数调制的高斯函数,是 $R^2 \rightarrow C$ 上的一个复值函数^[9]。

如今 Gabor 滤波器在计算机视觉的多个领域中已经被广泛应用。Gabor 滤波器具有许多优良的特性,如它可以对不同方向角 θ_i 下的图像特征进行提取等。

2.2 Quaternion-Gabor 滤波器

Thomas Bulow 在 1998 年就提出了相位匹配中的二维相位概念,并结合四元数将普通 Gabor 滤波器拓展成 Quaternion-Gabor 滤波器^[10]。改变了普通 Gabor 滤波器主要针对图像亮度信息进行滤波的状况。

Quaternion-Gabor 滤波器的实现如下。

普通二维 Gabor 滤波器公式重写如式(12):

$$h(x, y) = g(x, y) \exp[-2\pi i(Ux + Vy)] \quad (12)$$

其中 $g(x, y)$ 是高斯包络。将式(12)的指数部分进行欧拉展开:

$$\exp[2\pi i(Ux + Vy)] = \cos(2\pi Ux) \cos(2\pi Vy) - \sin(2\pi Ux) \sin(2\pi Vy) + i(\sin(2\pi Ux) \cos(2\pi Vy) + \cos(2\pi Ux) \sin(2\pi Vy)) \quad (13)$$

从式(13)中可以看出, Gabor 滤波器可以看成是由实部和虚部两个滤波器组成。这里用 $2\pi\mathbf{j}Vy$ 代替 $2\pi iVy$,有:

$$h_q = g(x, y) \exp(-2\pi iUx) \exp(-2\pi\mathbf{j}Vy) \quad (14)$$

对式(14)进行欧拉展开:

$$h_q = g(x, y) (\cos(2\pi Ux) \cos(2\pi Vy) - i \sin(2\pi Ux) \cos(2\pi Vy) - \mathbf{j} \cos(2\pi Ux) \sin(2\pi Vy) + k \sin(2\pi Ux) \sin(2\pi Vy)) \quad (15)$$

从式(15)可以看出, h_q 可以由一个实部和 i, j, k 三个虚部来表示,符合四元数的定义。这样,一个普通的二维 Gabor 滤波器就可以扩展到四元数域中。

在 Quaternion-Gabor 滤波器中,各个部分分别独立。在图像处理中,只要将灰度图输入并与 Gabor 滤波器的四个部分分别卷积,就可以实现滤波功能^[11]。经过 Gabor 滤波后的结果仍然是一个四元数。文献[12]介绍了利用四元小波进行立体图像对的相位匹配方法,本文将这种思想应用于视频前后帧的匹配中,以便获得场景的相对深度信息。

3 深度图的获取

利用运动估计获取深度信息是深度图获取技术中的一个非常重要手段。对于普通的场景,可以简单地分为前景和背景。而且前景以运动为主,而背景大多是静止的。如果能够从视频中分离出前景和背景,则深度信息可以利用式(16)得到:

$$d = k |m - m_{bg}| \quad (16)$$

其中: k 是常数,可以根据需要进行改变; m 是由运动估计获得的运动场的大小; m_{bg} 是通过建立背景模型获得的背景运动的大小。因此,获取视频深度图可以按图1所示的步骤进行。下面具体介绍各个步骤。

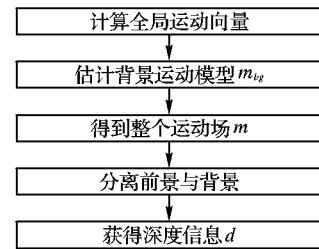


图1 Quaternion-Gabor 小波运动估计深度图获取流程

3.1 全局运动向量的计算

传统的运动向量计算方法有很多;其中利用亮度信息进行块匹配是最基本的方法;对已经编码过的 Mpeg 流进行解码也可以获得运动向量信息;在 H. 264 中通过寻找与当前图像最匹配参考图像的帧间预测技术也可以获取运动向量。本文采用基于 Quaternion-Gabor 小波的运动估计方法,利用 Quaternion-Gabor 小波系数的多方向多尺度空间进行运动向量计算。

3.2 背景运动模型的估计

为了能够拥有一个较大的深度来描述前景像素的运动,可以先估计出背景模型,然后再去除背景的运动。

假设背景是静止的,那么可以用式(17)来描述对应于无穷大深度的像素运动:

$$\begin{cases} m_x = \alpha \hat{y}_0 \frac{f_0}{f_1} + \beta f_1 + \left(\frac{f_1}{f_0} - 1\right) \hat{x}_0 + \beta \frac{\hat{x}_1 \hat{x}_0}{f_0} + \gamma \frac{\hat{y}_0 \hat{x}_1}{f_0} \\ m_y = -\alpha \hat{x}_0 \frac{f_0}{f_1} + \gamma f_1 + \left(\frac{f_1}{f_0} - 1\right) \hat{y}_0 + \beta \frac{\hat{x}_1 \hat{x}_0}{f_0} + \gamma \frac{\hat{y}_0 \hat{x}_1}{f_0} \end{cases} \quad (17)$$

其中:0 和 1 的下标分别表示连续的两帧序号; f 为焦距; $\hat{x} = x - x_c$ 和 $\hat{y} = y - y_c$ 分别为水平和垂直方向上像素点到光学中

心 (x_c, y_c) 的距离; α, β, γ 为摄像机的旋转角度; $\mathbf{m} = (m_x, m_y)$ 即是运动向量。

在采用摄像机拍摄的普通视频(单目或二维视频)中,可以把二次项忽略,式(17)可简化为式(18):

$$\begin{cases} m_x = \alpha \hat{y}_0 \frac{f_0}{f_1} + \beta f_1 + \left(\frac{f_1}{f_0} - 1 \right) \hat{x}_0 \\ m_y = -\alpha \hat{x}_0 \frac{f_0}{f_1} + \gamma f_1 + \left(\frac{f_1}{f_0} - 1 \right) \hat{y}_0 \end{cases} \quad (18)$$

其中:第一项为摄像机绕光轴的转动,第二项与摄像机平动有关,第三项由摄像机的聚焦远近决定。由于后面两项要比第一项更加常见,本文采用对背景模型做出平动-聚焦估计:

$$\begin{cases} m_x = p_x + s_x \hat{x}_0 \\ m_y = p_y + s_y \hat{y}_0 \end{cases} \quad (19)$$

其中:

$$\begin{cases} p_x = \frac{m_{x,1}x_0 - m_{x,0}x_1}{x_0 - x_1} \\ p_y = \frac{m_{y,1}y_0 - m_{y,0}y_1}{y_0 - y_1} \\ s_x = \frac{m_{x,0} - m_{x,1}}{x_0 - x_1} \\ s_y = \frac{m_{y,1} - m_{y,0}}{y_0 - y_1} \end{cases} \quad (20)$$

3.3 运动场的获得

根据 3.2 节中所提到的方法,由于前景物体一般位于画面中央,背景往往处于边界部分,选择图 2 所示靠近图像边界的像素点作为特征点是一种便捷而有效的方法。

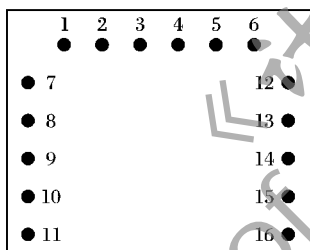


图 2 选择特征点

由式(19)可知,只要选择任意两个不处于同一列或同一排的两个点的运动向量就能根据式(20)得出相应参数。

这样,在图 2 所选择的特征点中就可以得到 8 组解 $A_i = (p_x, p_y, s_x, s_y), i = 1, 2, \dots, 8$ 。在最开始所选择的 16 个点中,由于只根据通常情况进行简单的判断,很容易出现偏差,这些点可能并不属于背景而属于前景。这种情况所获得的解会与其他解出现很大的误差,因此,要把这些点去除。本文采用相对简单且有效的计算方差的方法来排除异常点。

排除异常点后,将可靠点的算术平均值作为最终用来估计背景参数的特征块。将得到的参数结果代入式(19)~(20)即可以得到背景的运动场。

3.4 前景与背景的分离

经过在以上 3 个步骤,通过得到全局运动向量,可以估计出背景运动模型 m_{bg} 。经过对视频做局部运动估计,可以得到整个运动场 m 。这样就可以利用式(16)分离出前景与背景,将深度信息 d 计算出来。由于 k 是一个常数,用来表示最佳可视化深度信息。在一般情况下,可以将 k 值设为 1。

3.5 深度图的获取

图 3 是 skee 的原始图像,运动员从高处滑下,滑动速度快,运动剧烈。本文采用基于 Quaternion-Gabor 小波运动估计的深度图获取方法,利用 C++ 编程技术,将 skee 的深度图提取出来如图 4 所示。由图 4 可以看出该方法基本上将前景的轮廓都提取了出来,如运动员、雪橇、雪杖、被扬起的雪等运动的物体。而且整个深度图的变化很平滑,看上去感觉很舒服。



图 3 原始图像(skee)

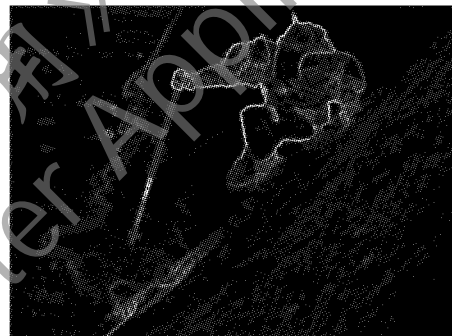


图 4 采用 Quaternion-Gabor 小波运动估计获得的深度图

图 5 是采用运动估计获得的深度图。由图 5 可以看出,雪杖黄色的上半截基本上看不出轮廓,雪橇和运动员左手的轮廓也不明显。深度图的整体感觉不平滑。

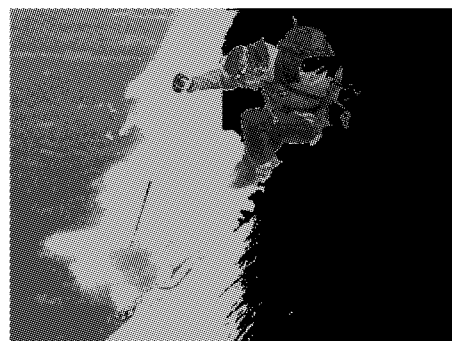


图 5 采用运动估计获得的深度图

4 结语

本文针对飞利浦公司“二维视频+深度图”格式的 3D 显示器,引入了最新用于图像处理的四元数,提出了利用 Quaternion-Gabor 小波运动估计获取深度图的方法。该方法通过对视频前后帧进行匹配,从而获得场景的相对深度信息,成功地提取出运动图像的深度图。与采用运动估计获得的深度图相比,该技术获得的深度图在整体和局部效果都要优越。该技术可以应用于“二维视频+深度图”格式 3D 显示器片源的制作,能够有效地解决 3D 行业影视内容匮乏的问题。

(下转第 244 页)

实时性,表1给出了本文与文献[5-6]的算法效率比较结果。需要注意的是,由于不同计算平台的硬盘 I/O 时间差异很大,本文的计算效率比较不包含 I/O 时间。

由于本文利用暗原色先验得到了较准确的大气光估计值,但是在实际情况中,当场景辐射与大气光相近或场景中存在大片白色区域时,暗原色先验失效,这就会对大气光的估计造成误差,使得复原结果发生颜色偏移;另一方面,虚拟无雾图像备选序列是通过对有雾图像的场景深度进行采样得到的,这就不可避免地带来原始信息的损失。可以通过增加图像序列长度 N 以及选择更优的接近人眼视觉的“像素质量评价标准”来充分挖掘出图像序列中所蕴含的原始信息来减小原始信息的损失。

表1 算法效率比较

方法	图像大小	
	800 × 457	600 × 400
本文方法	6.51	4.86
文献[5]方法	20.75	21.53
文献[6]方法	23.59	9.51

4 结语

为了解决雾天图像降质问题,本文提出了一种基于曝光融合算法的单幅图像去雾方法。利用暗原色先验对大气光值进行较高精度的估计,再结合雾天图像光学模型,在较宽范围内对场景深度进行采样,得到对应不同场景深度的图像序列,利用曝光融合算法,最终从雾的物理特性上有效恢复出清晰图像,且具有较好的实时性。本文算法同样也有一定的局限性,比如在暗原色先验失效的情况下,给出的复原图像会失真,因此有必要设计出适用范围更广的大气光估计方法。另外,场景深度的采样方式(本实验中采用等间隔均匀采样)和“像素质量评价标准”的不同选择都会对最终结果产生重要影响,上述内容都将是下一步工作的研究重点。

参考文献:

- [1] NARASIMHAN S G, NAYER S K. Contrast restoration of weather

degraded images [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(6): 713-724.

- [2] 陈功,王唐,周荷琴.基于物理模型的雾天图像复原新方法[J].中国图象图形学报,2008,13(5):888-893.
- [3] KOPF J, NEUBERT B, CHEN B, *et al.* Deep photo: Model-based photograph enhancement and viewing [J]. ACM Transactions on Graphics, 2008, 27(5): 116:1-116:10.
- [4] TAN R. Visibility in bad weather from a single image [C]// Proceedings of 2008 IEEE Conference on Computer Vision and Pattern Recognition. Washington, DC: IEEE Computer Society, 2008: 2347-2354.
- [5] HE KAIMING, SUN JIAN, TANG XIAOOU. Single image haze removal using dark channel prior [C]// Proceedings of 2009 IEEE Conference on Computer Vision and Pattern Recognition. Washington, DC: IEEE Computer Society, 2009: 1956-1963.
- [6] TAREL J-P, HAUTIERE N. Fast visibility restoration from a single color or gray level image [C]// Proceedings of 2009 IEEE International Conference on Computer Vision. Washington, DC: IEEE Computer Society, 2009: 2201-2208.
- [7] LEVIN A, LISCHINSKI D, WEISS Y. A closed form solution to natural image matting [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 30(2): 1-15.
- [8] SCHAUL L, FREDEMBACH C, SSSTRUNK S. Color image dehazing using the near-infrared [C]// Proceedings of 2009 IEEE International Conference on Image Processing. Washington, DC: IEEE Computer Society, 2009: 1629-1632.
- [9] MERTENS T, KAUTZ J, van REETH F. Exposure fusion: A simple and practical alternative to high dynamic range photography [J]. Computer Graphics Forum, 2009, 28(1): 161-171.
- [10] BURT P J, ADELSON E H. A multiresolution spline with application to image mosaics [J]. ACM Transactions on Graphics, 1983, 2(4): 217-236.
- [11] BURT P J, ADELSON E H. The Laplacian pyramid as a compact image code [J]. IEEE Transactions on Communication, 1983, 31(4): 532-540.

(上接第240页)

参考文献:

- [1] ZHANG LIANG, TAM W J. Stereoscopic image generation based on depth images for 3DTV [J]. IEEE Transactions on Broadcasting, 2005, 51(2): 191-199.
- [2] 周允.基于单目视频的深度信息预测算法的研究[D].兰州:兰州大学,2009.
- [3] FEHN C, PASTOOR R S. Interactive 3-DTV — Concepts and key technologies [J]. Proceeding of the IEEE, 2006, 94(3): 524-538.
- [4] 王琼华.3D显示技术与器件[M].北京:科学出版社,2011:154-179.
- [5] ALLEYSSON D, SIISSTRUNK S. Spatio-chromatic PCA of a mosaiced color image [C]// CGIV 2004: IS&T Second European Conference on Color in Graphics, Image, and Vision, LNCS 3195. Berlin: Springer-Verlag, 2004: 946-953.
- [6] MALONEK H. Quaternions in applied sciences a historical perspective of mathematical concept [C]// Proceedings of the 17th International Conference on the Application of Computer Science and Mathematics on Architecture and Civil Engineering. Weimar: [s. n.], 2003: 411-433.
- [7] SHI LILONG. Exploration in quaternion color [D]. Burnaby: Simon Fraser University, 2005.
- [8] ELL T A, SANGWINE S J. Hypercomplex fourier transforms of color images [J]. IEEE Transactions on Image Processing, 2007, 16(1): 22-35.
- [9] 玉振明,毛士艺,高飞.一种基于Gabor滤波的不同聚焦图像融合方法[J].航空学报,2005,26(2):219-223.
- [10] LU WEI, XU YI, YANG XIAOKANG, *et al.* Local quaternionic Gabor binary patterns for color face recognition [C]// ICASSP 2008: IEEE International Conference on Acoustics, Speech and Signal Processing. Washington, DC: IEEE Computer Society, 2008: 741-744.
- [11] 罗亚兰,陈锻生.基于彩色Gabor特征的人脸识别技术[J].中国图象图形学报,2008,13(2):242-246.
- [12] 徐奕.小波理论在相位匹配中的应用研究[D].上海:上海交通大学,2004.