



(12) 发明专利申请

(10) 申请公布号 CN 103561258 A

(43) 申请公布日 2014. 02. 05

(21) 申请号 201310442055. 3

(22) 申请日 2013. 09. 25

(71) 申请人 同济大学

地址 200092 上海市杨浦区四平路 1239 号

(72) 发明人 张冬冬 姚焜 刘典 陈艳毓

臧笛

(74) 专利代理机构 上海天协和诚知识产权代理

事务所 31216

代理人 叶凤

(51) Int. Cl.

H04N 15/00 (2006. 01)

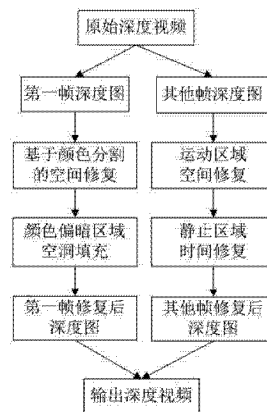
权利要求书1页 说明书5页 附图6页

(54) 发明名称

一种 Kinect 深度视频时空联合修复方法

(57) 摘要

一种 Kinect 深度视频时空联合修复方法。基于具有相似颜色的邻域像素应当具有相似的深度值这一假设,对于第一帧深度图像和其后所有帧深度图像中提取出的运动区域,使用对应彩色图像的颜色分割图来引导深度填充。考虑到一些没有有效深度值的颜色偏暗区域的存在可能导致上述方法的失败,先检测出颜色偏暗区域,然后用相同颜色偏暗区域内的有效深度值修复空洞区域。对于 Kinect 拍摄到的深度视频中的静止区域,如果当前帧深度图出现空洞像素,则用前一帧深度图的对应位置深度值填充。使用基于深度图像的绘制技术绘制虚拟视点,本发明提出的方法对得到的虚拟右视图的图像质量明显高于原始虚拟右视图的图像质量,可应用于 3D 绘制。



1. 一种 Kinect 深度视频时空联合修复方法,其特征在于,包括步骤有,

步骤 S1:对第一帧深度图,通过使用 mean-shift 算法对彩色图像进行颜色分割以获取对应的 label 图,接着对 label 图进行八联通域分割来进一步分离出各个连续的颜色块,当每一个颜色块内有效深度值所占比例超过一定阈值后,用该颜色块内所有有效深度值的中值填充空洞像素;

步骤 S2:对于经过步骤 S1 修复后的第一帧深度图,通过将对应原始彩色 RGB 图像转换为 YCbCr 图像,对 Y, Cb, Cr 三分量分别设置阈值,将满足所有阈值条件的区域提取出来作为颜色偏暗区域,并且将空间上相连的区域标定为同块区域,使用同块区域内所有有效深度值的中值填充该块内的空洞像素,最后使用膨胀腐蚀操作修复深度图像中细小空洞区域;

步骤 S3:对于第一帧之后的所有深度图像,通过对前后两帧彩图的灰度图进行差值处理,找出运动像素点,然后再将包含运动像素点的最小矩阵找出来,矩阵区域即视为运动区域,最后对运动区域进行步骤 S1 使用的空间修复;

步骤 S4:对于第一帧之后的所有深度图像,在抠出矩形运动区域后,剩余的部分即是背景静止区域;

若在静止区域出现空洞像素点,则用前一帧修复好的深度图的对应位置的深度值进行修复。

## 一种 Kinect 深度视频时空联合修复方法

### 技术领域

[0001] 本发明涉及图像 / 视频处理技术领域,可应用于 3D 绘制。

### 技术背景

[0002] 三维立体电视已经被许多人认为能带来更自然、更生活化的视觉娱乐体验。随着立体显示技术和视频处理技术的发展,3D 视频技术成为近年来的研究热点。目前,实现立体视频系统主要有两种方案:一种是多视方案,由多个摄像机阵列获取三维场景,在立体显示器上播放出来;另一种是“纹理+深度”方案,分别用彩色纹理视频和深度视频描述立体场景中的彩色纹理信息和深度信息,结合这两路视频信息,采用基于深度图像的绘制(DIBR)技术绘制虚拟视点,最后将合成的 3D 视频显示在显示器上。第二种方案相对于第一种方案,有数据传输带宽小、易于虚拟视点绘制等优点。

[0003] 要获取深度信息,主要有两种方法:一是采用立体匹配算法被动获取深度信息,这种方法很难在算法复杂度和获取深度视频的质量上取得折中。二是使用深度摄像头主动获取深度信息,目前深度摄像头主要分为两类,一类是 TOF(time of flight) 摄像头,另一类是 Kinect,它们都是通过发射和接收反射光或光斑来计算深度信息以实现深度信息的实时提取,也能够获取对应的实时彩色视频。Kinect 因为其价格低廉,并且能提取相对较高分辨率的深度图,吸引了更多的关注。但是 Kinect 提取的深度视频质量较差,在遮挡区域、光滑物体表面存在较大的深度信息缺失空洞,并且同一静止区域的深度值也会随时间变化而产生变化,因此必须进行填充修复处理。

[0004] 在已有的修复算法中,单纯的基于空间的修复算法比较多,对于深度视频中的光学噪声,一般采用中值滤波,双边滤波,联合双边滤波,线性插值等多种滤波方法进行噪声消减。例如文献 1(参见 Massimo Camplani and Luis Salgado. “Efficient Spatio-temporal Hole Filling Strategy for Kinect Depth Maps”, in Proceedings of SPIE, 82900E, 2012.) 中对于 Kinect 拍摄到的静态场景中空间域和时间域的相邻像素,迭代的使用一个联合双边滤波器去修复深度值。这些空间修复方法在消除细小空洞点,保护物体边界和平滑图像等方面效果显著。但是对于颜色偏暗区域或反光区域等容易产生大片空洞区域的填充修复,这些方法的效果并不明显。如果考虑到相似颜色块物体应当拥有相似深度值这点,填充修复大块空洞区域就可得到解决。另外受到光照影响, Kinect 拍摄到的深度视频中背景物体在不同时刻的深度值并不稳定,有时甚至会出现空洞点的现象,因此基于时间的修复方法也需要引进。目前已经出现了一些基于时间的修复算法,例如文献 2(参见 Sung-Yeol Kim, Ji-Ho Cho, Andreas Koschan, and Mongi A. Abidi, “Spatial and temporal enhancement of depth images captured by a time-of-flight depth sensor”, in Proc. of the IEEE International Conference on Pattern Recognition (ICPR), pp. 2358-2361, 2010.) 使用了一个联合双边滤波器和基于运动估计的时域连续来修复 TOF 摄像头获取的深度图像。这种方法能够有效地减少光学噪声和修复边界,但是却没有考虑反光和颜色偏暗区域的深度修复,而且这个方法的效果受到运动估计

准确性的影响。

## 发明内容

[0005] 针对大多数深度视频修复算法仅仅基于空间进行修复这个现象,本文提出了一种 Kinect 深度视频时空联合修复方法。基于具有相似颜色的邻域像素应当具有相似的深度值这一假设,对于第一帧深度图像和其后所有帧深度图像中提取出的运动区域,使用对应彩色图像的颜色分割图来引导深度填充。考虑到一些没有有效深度值的颜色偏暗区域的存在可能导致上述方法的失败,先检测出颜色偏暗区域,然后用相同颜色偏暗区域内的有效深度值修复空洞区域。对于 Kinect 拍摄到的深度视频中的静止区域,如果当前帧深度图出现空洞像素,则用前一帧深度图的对应位置深度值填充。概括地说,本发明给出技术方案实施步骤为:

[0006] 步骤 S1 :对第一帧深度图,通过使用 mean-shift 算法对彩色图像进行颜色分割以获取对应的 label 图,接着对 label 图进行八联通域分割来进一步分离出各个连续的颜色块,当每一个颜色块内有效深度值所占比例超过一定阈值后,用该颜色块内所有有效深度值的中值填充空洞像素。

[0007] 步骤 S2 :对于经过步骤 S1 修复后的第一帧深度图,通过将对应原始彩色 RGB 图像转换为 YCbCr 图像,对 Y,Cb,Cr 三分量分别设置阈值,将满足所有阈值条件的区域提取出来作为颜色偏暗区域,并且将空间上相连的区域标定为同块区域,使用同块区域内所有有效深度值的中值填充该块内的空洞像素,最后使用膨胀腐蚀操作修复深度图像中细小空洞区域。

[0008] 步骤 S3 :对于第一帧之后的所有深度图像,通过对前后两帧彩图的灰度图进行差值处理,找出运动像素点,然后再将包含运动像素点的最小矩阵找出来,矩阵区域即视为运动区域,最后对运动区域进行步骤 S1 使用的空间修复。

[0009] 步骤 S4 :对于第一帧之后的所有深度图像,在抠出矩形运动区域后,剩余的部分即是背景静止区域。若在静止区域出现空洞像素点,则用前一帧修复好的深度图的对应位置的深度值进行修复。

[0010] 区别于现有技术,上述技术方案体现出具有技术贡献的关键技术要点:

[0011] 1、针对大部分 Kinect 深度图像修复算法没有考虑对于颜色偏暗区域或反光区域等容易产生大片空洞区域的填充修复这个问题,本发明提出了颜色偏暗区域的空洞修复方法。首先要检测出颜色偏暗区域,通过将原始彩色 RGB 图像转换为 YCbCr 图像,对 Y,Cb,Cr 三分量分别设置阈值,将满足所有阈值条件的区域提取出来作为颜色偏暗区域。然后将空间上相连的区域标定为同块区域,使用同块区域内所有有效深度值的中值填充该块内的空洞像素。

[0012] 2、本发明结合了空间修复算法和基于运动区域检测的时间修复算法,以解决单纯使用空间修复算法可能会出现修复不准确和修复时间过长的的问题,使得修复效果更优。

[0013] 3、通过对颜色分割后得到的 label 图进一步进行八联通域分割来分离不连续分布的颜色块,使得颜色分割结果更准确。

[0014] 4、在基于时间域的修复算法中,对检测出的运动点,用最小的矩阵区域包含它们,将矩阵区域视为运动区域,方便运动区域和静止区域的区分修复。

[0015] 由此,与现有技术相比之,本发明方法技术方案具有的有益效果:颜色偏暗区域的空洞修复效果明显,在保证同样拍摄场景的前提下,使用本发明提出的 Kinect 深度视频修复方法得到的深度图像对比于原始深度图像,修复效果很好。使用基于深度图像的绘制 (DIBR) 技术绘制虚拟视点,本发明提出的方法对应得到的虚拟右视图的图像质量也明显高于原始虚拟右视图的图像质量。

## 附图说明

[0016] 图 1 是本发明一种 Kinect 深度视频时空联合修复方法流程图。

[0017] 图 2 是本发明实例 Kinect 拍摄到的三帧连续的彩色图像和深度图像。

[0018] 图 3 是本发明实例 Kinect 拍摄到的第一帧经过步骤 S1 得到的 (a) 彩色图像颜色分割图和 (b) 修复后深度图。

[0019] 图 4 是本发明实例 Kinect 拍摄到的第一帧经过步骤 S2 得到的 (a) 颜色偏暗区域 (不同的颜色代表了不同的颜色偏暗区域,灰色区域不是偏暗区域) 和 (b) 修复后深度图。

[0020] 图 5 是本发明实例经过步骤 S3 得到的 (a) 第一帧灰度图, (b) 第二帧灰度图, (c) 第二帧的运动点分布图, (d) 第二帧彩色图中放大后的运动区域子图, (e) 第二帧运动区域对应的深度子图和 (f) 运动区域经过修复后的深度子图。

[0021] 图 6 是本发明实例最终的效果图 (a) 未处理的深度图像, (b) 修复后的深度图像, (c) 未进行深度修复得到的虚拟右视图, (d) 进行深度修复得到的虚拟右视图, (e) 放大的 (c) 中的显示器区域, (f) 放大的 (d) 中的显示器区域, (g) 放大的 (c) 中的椅子右腿区域, (h) 放大的 (d) 中的椅子右腿区域。

## 具体实施方式

[0022] 下面以具体实例结合附图对本发明作进一步说明:

[0023] 本发明使用由 Kinect 拍摄到的一段深度视频序列和对应的彩色视频序列来检测提出方法的效果。视频序列由 100 帧图像组成。图像分辨率是  $640 \times 480$ 。所有涉及实例采用 MATLAB7 作为仿真实验平台。

[0024] 本发明的流程图如图 1 所示:对于第一帧深度图,只使用空间修复。使用其对应彩色图像的颜色分割图来引导初始的深度填充,然后对颜色偏暗区域进行空洞修复以进一步提高深度图质量。对于第一帧之后的所有深度图,首先提取出运动区域,单独对运动区域进行空间修复,然后用前一帧修复好的深度图对应位置的深度值填充余下静止区域内的空洞点。下面结合每个步骤详细描述本实例:

[0025] 步骤 (1),对当前需要修复的深度图像,这里选定实例图 2 中的第一帧深度图 (d),先对其对应的彩色图像,即实例图 2 中的第一帧彩色图 (a) 进行 mean-shift 颜色分割,得到标识颜色信息的 label 图和颜色分割图,即实例图 3(a)。设图像分辨率为  $n \times m$ ,每个像素点  $(i, j)$  都有一个代表颜色块索引的标签  $L_{ij}$ ,  $x_{ij}$  是该点的彩色值。对所有颜色数据聚类后,满足空间距离小于  $h_c$  且颜色空间距离小于  $h_l$  的一些点会被聚类。 $C_p$  代表特征集,  $q$  是所有特征个数。 $g(x)$  是图像特征空间中核函数的负导数。那么图像分割后得到的 label 图可用下式描述:

$$[0026] \quad \left\{ \begin{array}{l} L_{ij} = \{p | z_{ij} \in C_p\} \quad p = 1, 2, \dots, q \\ z_{ij} = \lim_{k \rightarrow \infty} \frac{\sum_{j=1}^m \sum_{i=1}^n x_{ij} g \left( \left\| \frac{y_k^c - x_{ij}^c}{h_c} \right\|^2 \right) g \left( \left\| \frac{y_k^l - x_{ij}^l}{h_l} \right\|^2 \right)}{\sum_{j=1}^m \sum_{i=1}^n g \left( \left\| \frac{y_k^c - x_{ij}^c}{h_c} \right\|^2 \right) g \left( \left\| \frac{y_k^l - x_{ij}^l}{h_l} \right\|^2 \right)} \quad y_1 = x_{ij} \end{array} \right. \quad (1)$$

[0027] 考虑到一些具有相同颜色标签的区域并不相连,使用八联通域分割来进一步分离出各个连续的颜色块。设  $D_{ij}^k$  为第  $k$  个颜色块内空洞点  $(i, j)$  经过修复后的深度值,  $S^k$  是彩色图像中第  $k$  个颜色块,  $D(S^k)$  代表块内所有的有效深度值,  $n(S^k)$  是  $S^k$  中所有像素点个数,  $n(D(S^k))$  是  $S^k$  中拥有有效深度值的像素点个数,  $T$  是一个阈值。接下来就使用下式来填充空洞:

$$[0028] \quad D_{ij}^k = \begin{cases} \text{median}\{D(S^k)\} & \text{if } \frac{n(D(S^k))}{n(S^k)} \geq T \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

[0029] 经过步骤 (1) 得到的修复后深度图如图 3(b) 所示。

[0030] 步骤 (2), 经过步骤 (1) 修复后的深度图像仍可能存在大块的隶属于颜色偏暗区域的空洞点, 为此首先将当前帧的原始彩色 RGB 图像转换为 YCbCr 图像, 对 Y, Cb, Cr 三分量分别设置阈值, 将满足所有阈值条件的区域提取出来作为颜色偏暗区域。设  $luma(i, j)$  为点  $(i, j)$  的 Y 分量值,  $blue\_diff(i, j)$  为蓝色色度差值,  $red\_diff(i, j)$  为红色色度差值,  $T_{luma}$ ,  $T_{blue\_diff}$  和  $T_{red\_diff}$  都是用大津法计算出来的阈值。1 是设置为 0.8 的参数。采用下式后具有非零亮度值的区域就是颜色偏暗区域, 图 4(a) 是从事例彩色图像图 2(a) 中提取出的颜色偏暗区域。

$$[0031] \quad luma(i, j) = \begin{cases} luma(i, j) & \text{if } \begin{cases} luma(i, j) \leq \lambda T_{luma} \\ blue\_diff(i, j) \leq T_{blue\_diff} \\ red\_diff(i, j) \leq T_{red\_diff} \end{cases} \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

[0032] 找到偏暗区域后, 将空间上相连的区域标定为同块区域, 使用同块区域内所有有效深度值  $D_{xy}$  的中值  $\tilde{D}_y$  填充该块内的空洞像素  $(i, j)$ , 如式 4。最后使用半径为 1, 高度为 3 的椭圆结构元素的膨胀腐蚀操作 (分别由  $\oplus$  and  $\odot$  表示) 修复深度图像中细小空洞区域  $D_{ij}$ , 修复后深度值为  $\tilde{D}_y$ , 如式 5。  $\oplus$  and  $\odot$  分别表示膨胀操作和腐蚀操作

$$[0033] \quad \tilde{D}_y = \text{median} \{Y D_{xy}\} \quad (x, y) \in \text{the same dark region} \quad (4)$$

[0034]

$$\tilde{D}_y = (D_{ij} \oplus b) \odot b \quad (5)$$

[0035] 经过步骤 (2) 得到的修复后深度图如图 4(b) 所示。

[0036] 步骤 (3), 对于第一帧之后的所有深度图, 如式 6 所示, 通过对前后两帧彩图的灰度图 (图 5(a) 和 (b)) 进行差值处理, 用  $\text{difference}(i, j)$  表示点  $(i, j)$  的灰度差值, 找出非零的  $\text{motion\_mask}(i, j)$  对应的所有运动像素点  $(i, j)$ , 如图 5(c) 所示, 再将包含运动像素点的最小矩阵找出来, 矩阵区域即视为运动区域, 图 5(d) 和 (e) 即分别是运动区域彩色

子图和运动区域深度子图,最后对运动区域进行步骤(1)使用的空间修复,得到的修复后深度子图如图5(f)所示。

$$[0037] \quad \text{motion\_mask}(i, j) = \begin{cases} 1 & \text{if } \text{difference}(i, j) \geq T_a \\ 0 & \text{Otherwise} \end{cases} \quad (6)$$

[0038] 步骤(4),对于第一帧之后的所有深度图像,在抠出矩形运动区域后,剩余的部分即是背景静止区域。若在静止区域出现空洞像素点 $(i, j)$ ,则用前一帧(第 $k-1$ 帧)修复好的深度图的对应位置的深度值 $\tilde{D}_{ij}^{k-1}$ 进行修复,如下式所示:

$$[0039] \quad \tilde{D}_{ij}^k = \tilde{D}_{ij}^{k-1} \quad (i, j) \in \text{static area} \quad (7)$$

[0040] 对于每一帧深度图而言,其都由修复后的运动区域深度图和静止区域深度图组成。

[0041] 图6(a)和(b)分别显示了第三帧未处理的深度图和使用本发明处理后的深度图。使用DIBR技术绘制得到的对应右视图显示在图6(c)和(d)中。

[0042] 从这些图像中,我们能够看出来本发明能够有效地修复深度图像,进而为3D绘制提供更清晰的深度图。用修复后的深度图像绘制的虚拟视图会比用原始Kinect拍摄到的深度图像绘制的虚拟视图具有更好的视觉质量。特别是一些比较重要的细节区域,如女孩的头发,显示器和椅子腿,修复后的深度图能够为3D绘制提供更平滑更好的纹理。

[0043] 本发明的创新点:

[0044] ➤采用了目前较少使用的时空联合的修复方法对由Kinect拍摄到的原始深度视频进行填充修复,提高了修复效率,效果显著。

[0045] ➤提出了图像中颜色偏暗区域的查找方法,并对深度图中偏暗区域的空洞像素点进行了填充,取得了不错的效果。

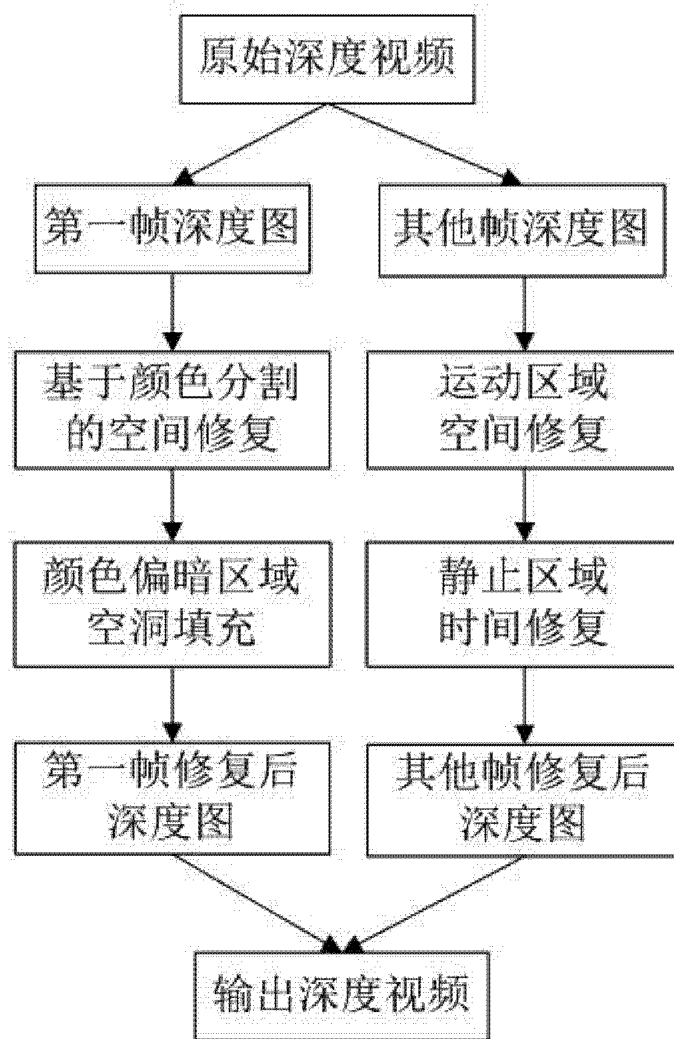


图 1



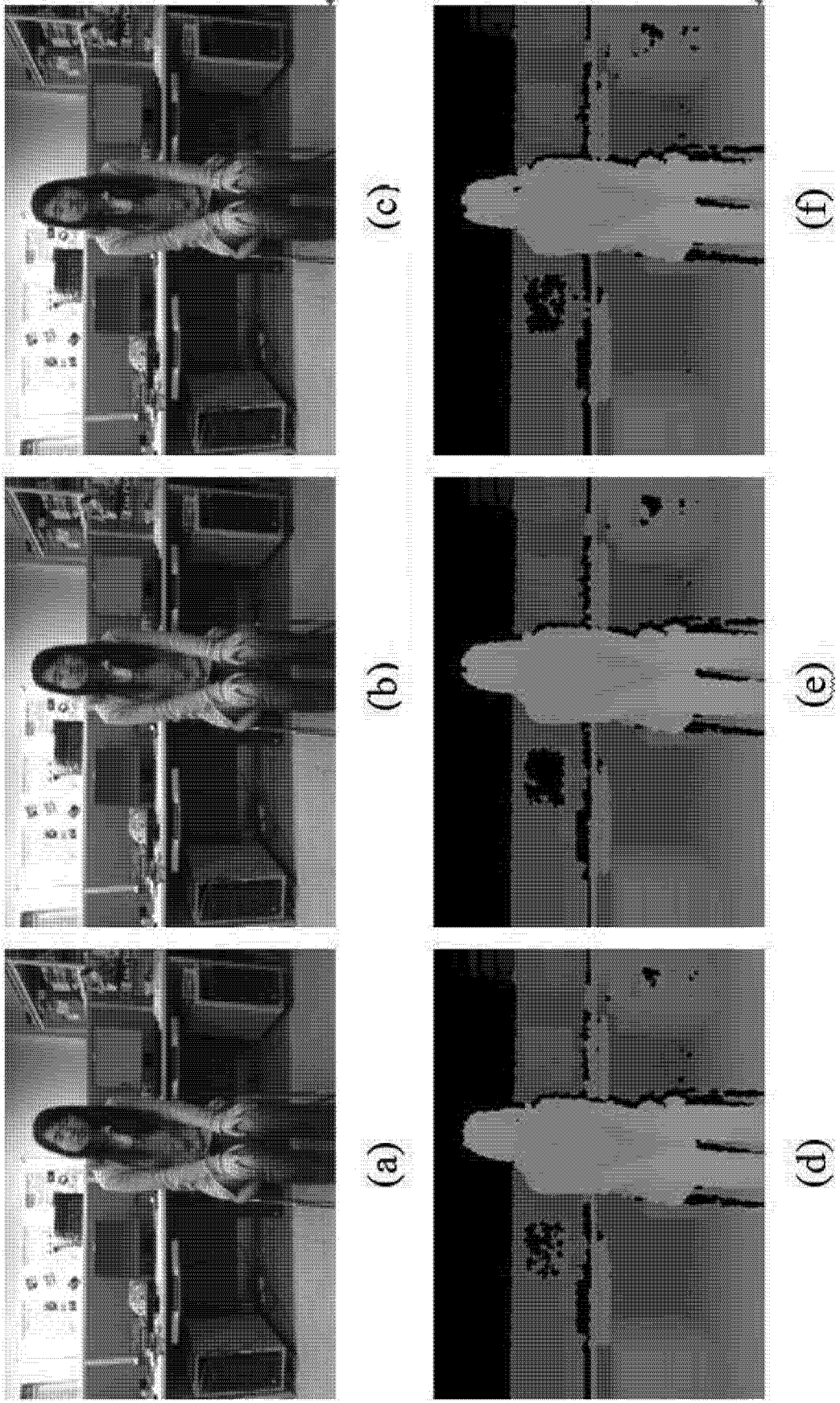


图 2

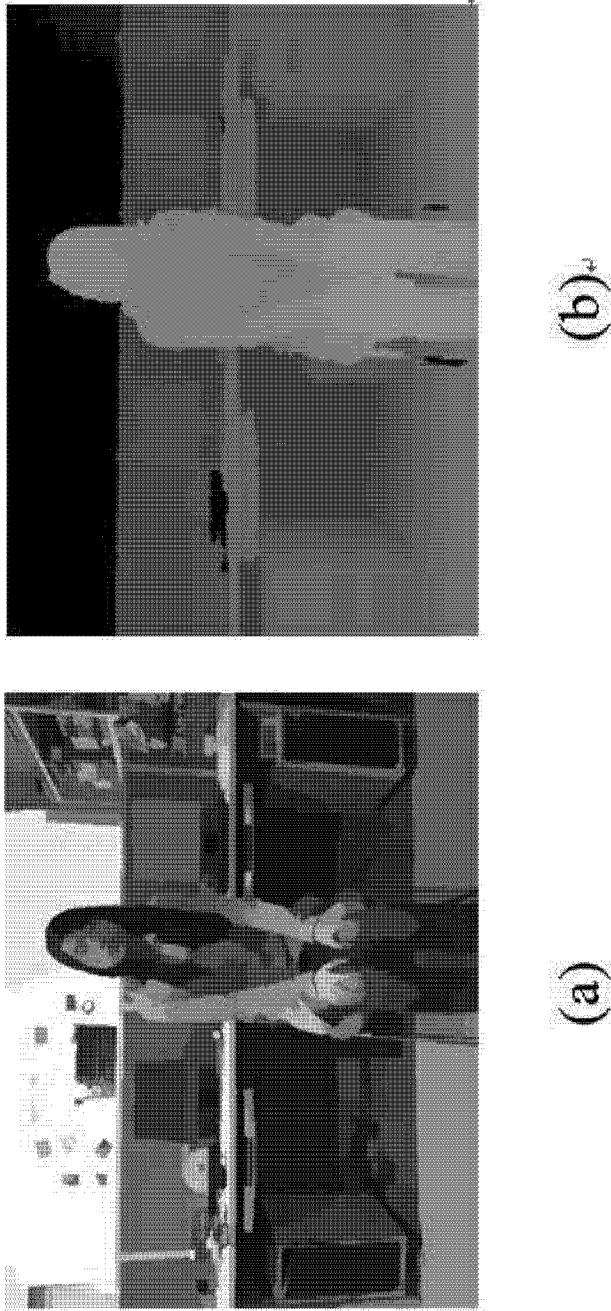


图 3

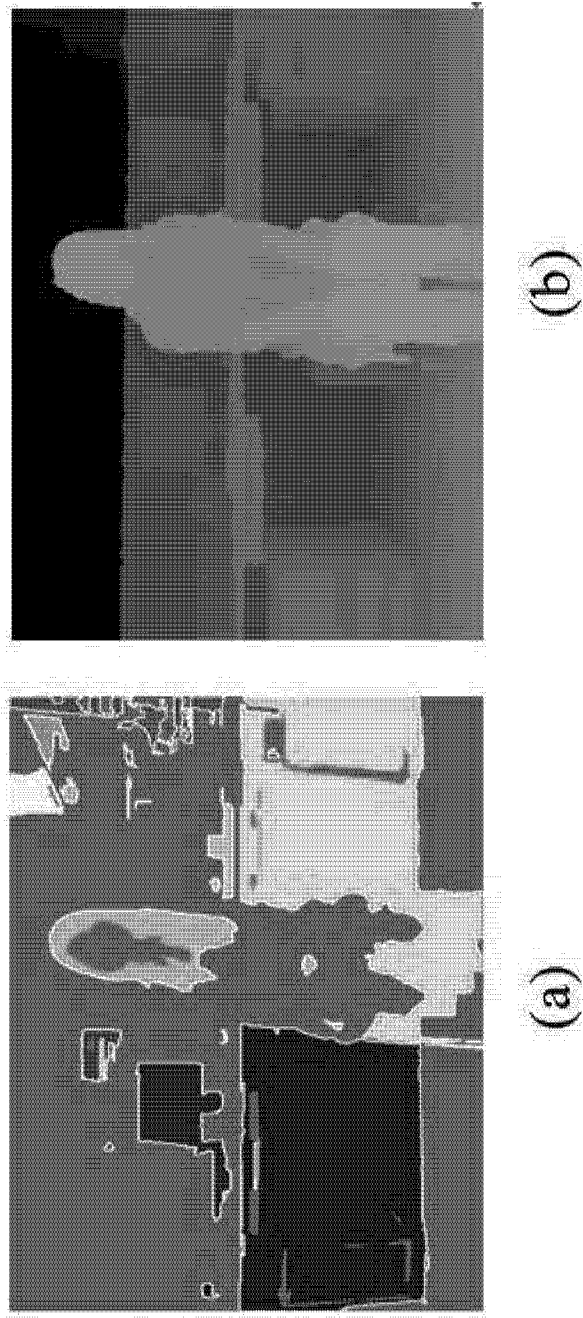


图 4

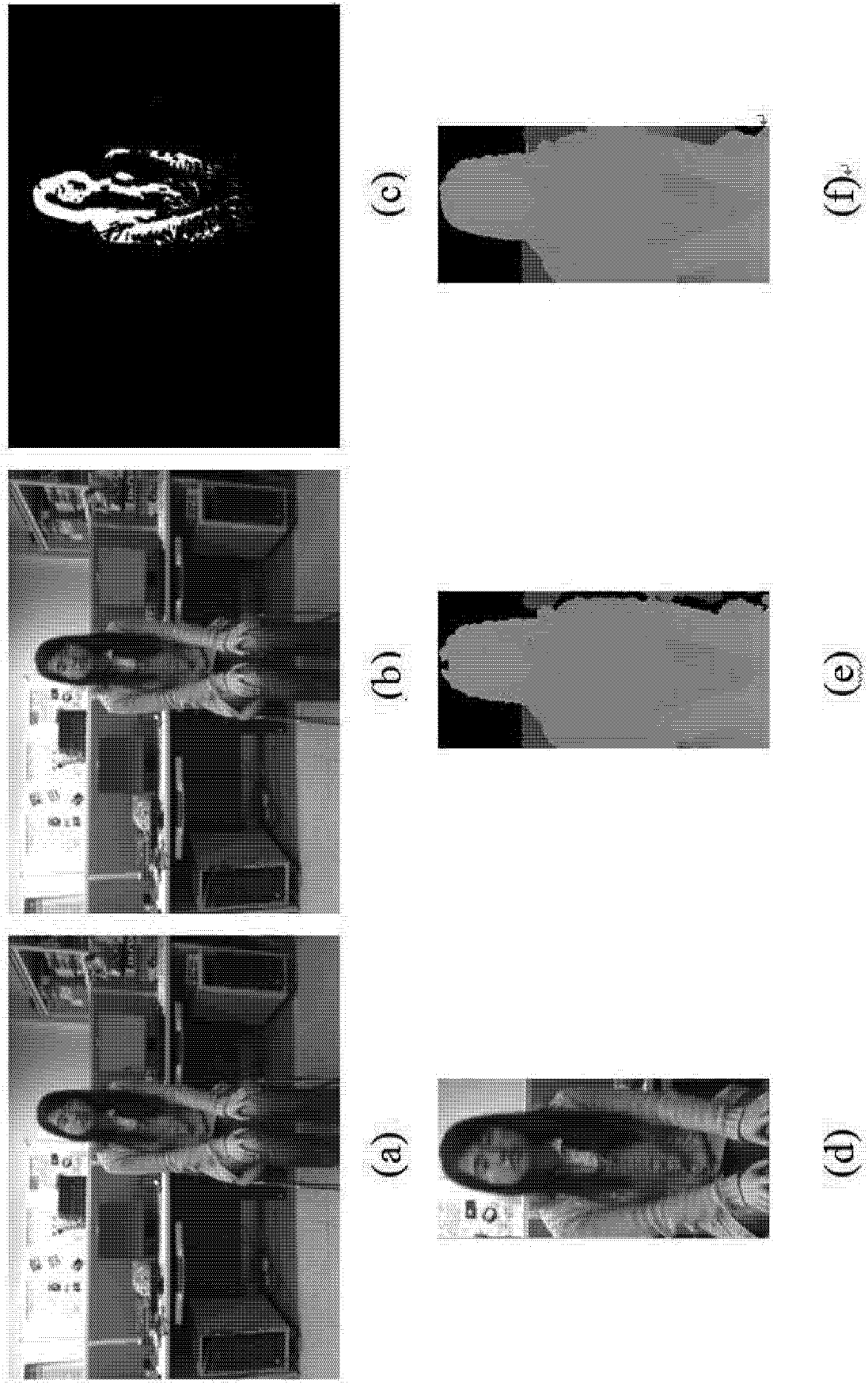


图 5

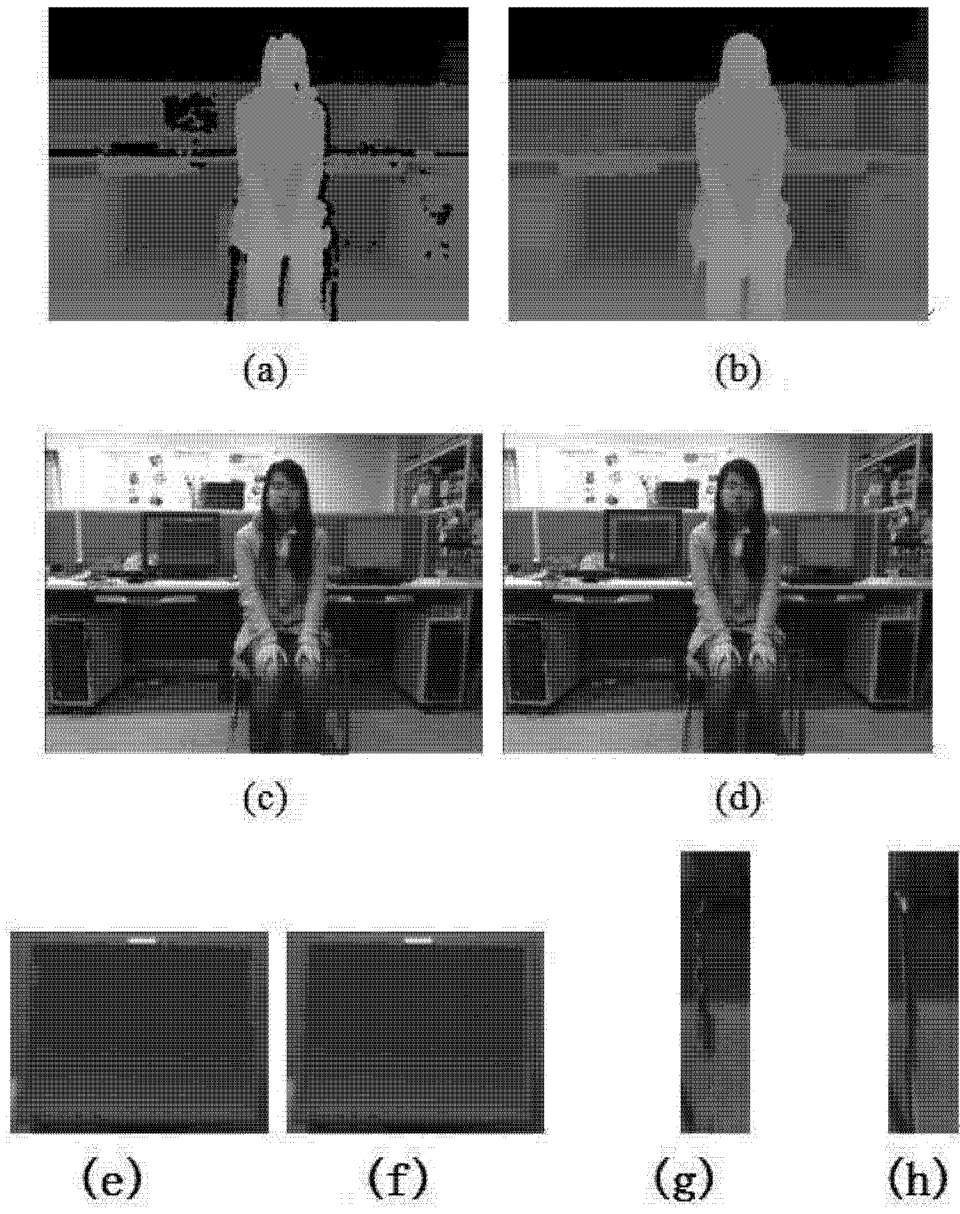


图 6