

# 基于背景复原的视频目标跟踪研究

卢永祥 赖文豪

(武夷学院 计算机教学实验中心,福建 武夷山 354300)

**摘 要:**通过对视频图像中只有一个运动目标、以及含有多目标的视频进行测试,实验研究结果表明:小波神经网络将小波变换良好的时频局域化特性和神经网络的自学习功能相结合,具有最佳的函数逼近能力和容错能力。

**关键词:**视频;目标跟踪;小波神经网络

**中图分类号:**TP391.41 **文献标识码:**A **文章编号:**1674-2109(2014)05-0061-03

究得出实验结论。

## 引言

视频目标跟踪是计算机视觉研究领域的一个重要组成部分,它融合了计算机图像处理、模式识别、人工智能以及自动控制等诸多相关领域的知识,形成了一种能从图像序列中自动检测目标,提取目标位置信息,自动跟踪目标的技术<sup>[1]</sup>。

随着信息科学技术的快速发展,对视频目标跟踪定位提出更高的要求。在视频目标发生旋转、被遮挡等情况下,如何通过视频图像实现对所需要的视频目标进行快速并且准确的定位成为广大学者的研究方向。国内外许多学者对视频目标跟踪做出了大量科学研究,提出了许多的视频目标跟踪算法。其中,神经网络算法可以用来解决传统视频目标跟踪方法中难以解决的需要确定机动目标的数学模型问题,提出基于小波神经网络的视频目标跟踪方法来提高视频目标跟踪的估计精度。本文从视频图像背景的更新,视频运动目标的移除,视频运动目标中关键区域检测进行研

## 1 图片背景的更新

通常要在视频流中获得实时的背景图像是很难的。但是,为了获取运动目标的活动区域,提取背景是必要的一步。一种较好的背景提取方法,是使用的灰度图像来实现<sup>[2]</sup>。其主要步骤如下:

(1) 为了提取运动的目标,计算相邻两对帧之间的相减图像,即帧  $t-1$  和  $t$  之的减图像,以及帧  $t$  和  $t+1$  之间的减图像,并分别将这两个图像定义为  $I^{t-1,t}$  和  $I^{t,t+1}$ 。接下来选择一个合适的阈值将这两个图像进行二值化,并且运用中值滤波技术对这两个差分图像进行处理,去除其中的噪音,并对其进行平滑。接下来,使用数学形态学中的膨胀运算和闭运算对图像进行处理,使运动的目标区域更加的明显,其中结构元素的定义是在实验中具体得出。

(2) 接下来,分别定义  $I^{t-1}, I^t, I^{t+1}$  为帧  $t-1, t, t+1$  的图像。运用  $SB^{t,t-1} = I^{t-1} - I^{t,t-1}, SB^{t,t+1} = I^t - I^{t,t+1}$  公式进行减运算,接下来对得到的后续图像进行逻辑与。

(3) 重复步骤 1 和 2 的过程,直到恢复的图像达到一定的标准为止,通常我们用恢复到原图像的 95% 为阈值,这是一个迭代的过程。实验结果图像如下图 1 所示。

收稿日期:2014-08-29

作者简介:卢永祥(1985-),男,汉族,助理实验师,主要研究方向:网络工程。

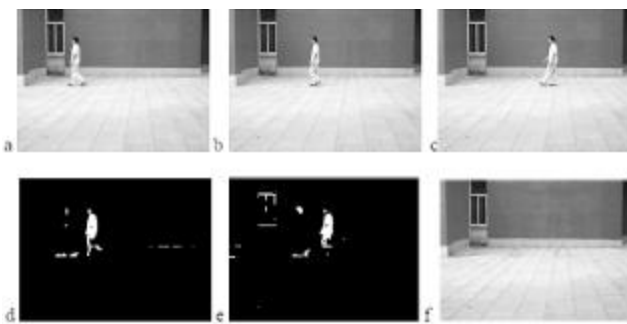


图 1 背景恢复

Fig.1 Background Restoration

## 2 运动目标的移除

得到了背景图像以后,可以提取出运动的目标,操作如下:

(1)定义背景图像为  $B$ ,并且定义计算后的差分图像为  $SIB^i = I^i - B$ 。

(2)通过合适的阈值,将转换为二值图像,并运用中值滤波去掉这个差分图像的噪音并使它平滑。

(3)运用数学形态学中的闭运算,选定合适的结构元素对图像进行处理。

经过这些步骤以后,可以获得一些连通的区域;再通过计算,找到一个连通面积最大的区域,使用一个最大的外接矩形来标识这个面积最大的区域。这样,可以去掉一些面积较小的冗余区域。最终图像如下图 2 所示。



图 2 定位过程

Fig.2 Positioning Process

## 3 运动目标中关键区域检测

在一些视频流中,运动目标的关键区域(例如人面部)的几何特点并不是十分的明显,我们只得使用颜色信息来完成脸部信息的提取<sup>[3]</sup>。

在运动目标的关键区域检测中,表面的颜色是一

个十分重要的信息,它一般只和亮度有关,而且它独立于面部的其它的一些特征,而且对于表情的变换和旋转保持了不变的特点。就一般而言,关键区域(例如人脸)的颜色和其他部位颜色有很大的差别。因此,在我们的问题中,使用颜色信息来寻找关键区域(例如人脸)是一个有效的方法。

这样就需要一个合适的颜色空间来对关键区域(例如人脸)的颜色建模。在通常的方法中,所使用的 RGB 的空间通常用于存储图像。这个空间反映了三个维的信息。但是在皮肤的颜色中亮度和色度是十分重要的,所以这个颜色空间不适合我们的问题。经过研究发现, YCrCb 空间更符合人的视觉特征,而且也广泛的应用于视频压缩标准。(比如 MPEG 和 JPEG)。颜色空间的转换的方法如下式所示。

$$Y = 0.2990r + 0.5870g + 0.1140b$$

$$Cb = -0.1787r - 0.3313g + 0.5000b + 128$$

$$Cr = 0.500r - 0.1487g - 0.0813b + 128$$

其中  $Y$  是亮度信号,  $CrCb$  是色度信号,亮度对于色度的影响很小。 $Cr$  和  $Cb$  在二维空间中保持相对独立。这两个方面可以限制颜色空间的分布区域,所以使用 YcrCb 空间。在这个过程中,选用了不同亮度、环境下的 55 幅包括关键区域(例如人脸)部区域的图像,并通过  $Cr$  和  $Cb$  来取得关键区域(例如人脸)的区域。所选的图像统一规格,大小均是  $103 \times 106$  像素。关键区域(例如人脸)颜色的分布区域范围如图 3 所示。

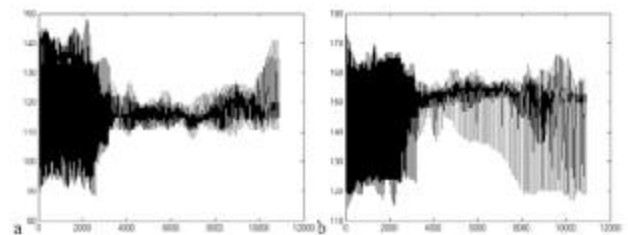


图 3 关键区域颜色的分布范围

Fig.3 the range of the color of the key area

在上面这两副图像里,其中  $X$  轴表示 55 幅图像总的像素数。 $Y$  轴分别表示  $Cr$  和  $Cb$  的值。通过结果图像找到皮肤颜色的像素的值比较集中的一个区域,其中有 97% 的皮肤像素的颜色都集中以下的范围内  $cb(i,j) \in [93, 133], cr(i,j) \in [123, 165]$ : 运用皮肤颜色的区域,我们可以把关键区域(例如人脸)变成二值图

像。方法如下所示:

$$M(i,j)=\begin{cases} 1, & \text{cb}(i,j) \in [193, 133], \text{cr} \in [123, 165] \\ 0, & \text{other} \end{cases}$$

经过二值化以后,图像仍还有一些小的区域,这些小区域一般是手臂等颜色和皮肤颜类似的部分。在后续工作中,可以通过数学形态学的闭运算的方法以及中值滤波技术将它们移除掉。结果如下图 4 所示:

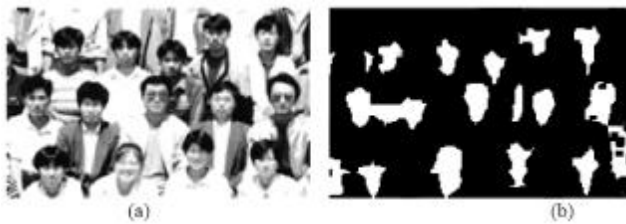


图 4 关键区域(例如人脸)区域的提取

Fig.4 The area extraction of the key area (such as face)

#### 4 实验结果和结论

在实验中,我们使用的是在室内和室外拍摄的视频图像。跟踪目标为视频中的运动目标,室内的图像都是在实拍的,包含有墙和窗户。室外的图像是在白天拍摄的,里面含有树和草。所有的视频剪辑都是在摄影机上进行,没有任何放大缩小或平移;它们都通过 XVID 格式压缩,图像大小为 640×480 像素。对选取的 100 幅图像,使用不包含关键区域(例如人脸)区域的人的图像进行小波神经网络的训练和学习。在实验过程中,对只有运动中一个目标的、以及含有多目标的视频都进行了测试。图 6 显示了对一个目标的跟踪情况;图 7 显示了对多个目标的跟踪画面。研究表明:小波神经网络将小波变换良好的时频局域化特性和神经网络的自学习功能相结合,因而具有最佳的函数逼近能力和容错能力,用此种方法与其他预测方法得出的结果进行比较,图 5、6 显示了更为逼真的学习预测效果。

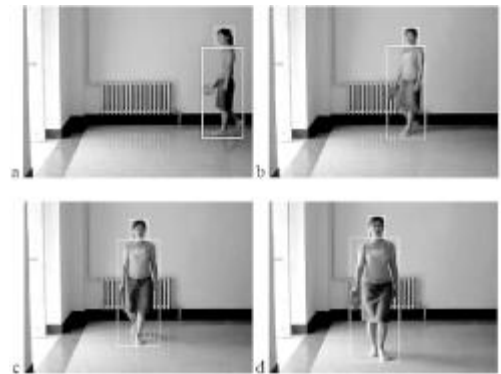


图 5 只含一个人的跟踪

Fig.5 Track Containing Only One Person

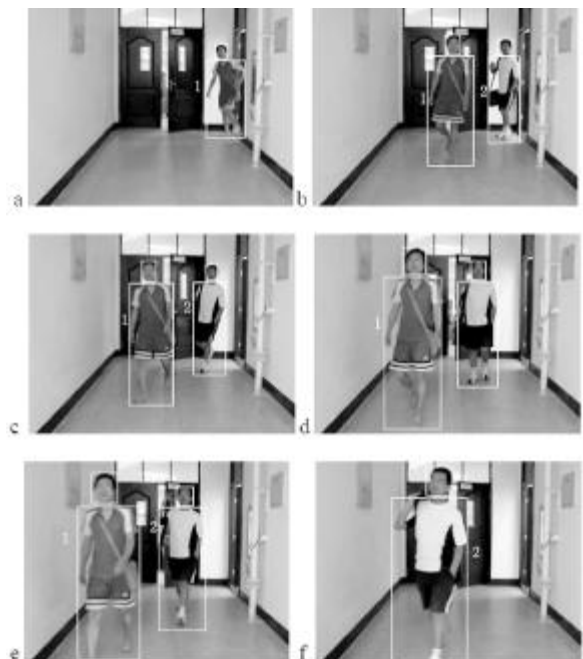


图 6 含多个人的跟踪

Fig.6 Track Containing More Than One Person

#### 参考文献:

- [1] 孙鑫.基于水平集的目标轮廓跟踪算法的研究[D].哈尔滨:哈尔滨工业大学,2010:4-5.
- [2] 罗三定.一种医学图像的轮廓提取方法[J].计算机工程,2010(5):218-219.
- [3] 徐全生,甄颖.基于知识库进行人脸图像特征提取的研究[J].沈阳工业大学,2000,22(1):43-46.

(下转第 76 页)