

基于文字特征的增强现实识别注册方法

李雪婷^{1,2*}, 党建武^{1,2}, 王阳萍^{1,2}, 高凡一^{1,2}

¹兰州交通大学电子与信息工程学院, 甘肃 兰州 730070;

²甘肃省人工智能与图形图像处理工程研究中心, 甘肃 兰州 730070

摘要 针对增强现实识别注册的精度易受纹理影响且缺少以文字为识别目标的问题, 提出了一种基于文字特征的识别注册方法。在模板图像处理中, 将降采样与 2 幂次采样相结合构建多尺度金字塔, 实现尺度不变性; 采用基于快速视网膜特征改进的算法提取文字特征点, 实现基于文字特征的增强现实系统。实验结果表明: 所提算法能较准确地提取图像文字特征点, 减弱了纹理对精度的影响, 适用于对文本图像的识别注册。基于此方法实现的增强现实系统在不同尺度角度和目标被部分遮挡的情况下仍可以较准确地进行识别注册。

关键词 机器视觉; 增强现实; 识别注册; 文字特征; 特征提取

中图分类号 TP391 文献标志码 A

doi: 10.3788/LOP57.021502

Augmented Reality Recognition Registration Method Based on Text Features

Li Xueting^{1,2*}, Dang Jianwu^{1,2}, Wang Yangping^{1,2}, Gao Fanyi^{1,2}

¹ School of Electronic and Information Engineering, Lanzhou Jiaotong University, Lanzhou, Gansu 730070, China;

² Gansu Provincial Engineering Research Center for Artificial Intelligence and Graphics & Image Processing, Lanzhou, Gansu 730070, China

Abstract A text feature-based recognition registration method is proposed to address the problems that the accuracy of the augmented reality recognition registration is prone to be affected by textures and that there is lack of text-based recognition targets. In template imaging processing, two sampling methods, i. e., downsampling and 2 power-based sampling, are combined to construct a multi-scale pyramid, achieving scale invariance. The text feature points are extracted based on an improved algorithm that uses fast retina keypoint (FREAK). Finally, the augmented reality system based on text features is realized. The experimental results show that the proposed algorithm can extract text feature points accurately and reduce the effect of the texture on accuracy, and it is applicable to the recognition registration for text images. Further, the augmented reality system based on the improved method can realize recognition registration in the situation that the target is partially obscured under different scales.

Key words machine vision; augmented reality; recognition registration; text features; feature extraction

OCIS codes 150.0155; 100.3008; 100.5010

1 引言

增强现实(AR)^[1]是虚拟现实(VR)的一项重要分支技术,其目的是将由计算机生成的虚拟模型实时、准确地叠加于真实场景中,实现虚、实场景的融合,达到“实中有虚”的呈现效果^[2]。AR具有虚实

融合、实时交互和 3 维注册 3 个技术特征,已被广泛应用于教育、医疗、娱乐等领域^[3]。

AR 系统主要包括场景视频获取、图像识别、3 维注册、虚实融合等模块,其中,图像识别和 3 维注册是最核心的研究内容^[4]。现有的识别注册方法大多基于计算机视觉,主要是对自然场景中的待识别

收稿日期: 2019-05-06; 修回日期: 2019-06-11; 录用日期: 2019-06-24

基金项目: 甘肃省科技计划(18JR3RA104)、甘肃省教育厅科技项目(2017D-08)

* E-mail: LiXueting510@163.com

目标实现增强^[5],主要分为两大类:基于标识的识别注册方法和基于自然特征(即无标识)的识别注册方法^[6-7]。其中,基于标识的识别注册方法已日趋成熟,但在真实环境中放置人工标志会带来视觉污染,且存在稳健性差、无法解决遮挡等缺点,因而在应用中受到了一定限制^[8]。近年来,基于自然特征的识别注册方法正日益受到关注,通过对自然场景中待识别的内容进行获取、理解以及计算虚实注册矩阵,实现了虚实融合,克服了基于标识的识别注册方法中存在的问题^[9]。其中,待识别内容的获取与理解以特征点匹配策略为基础,而图像匹配成功的关键在于特征检测。在特征检测算法中,加速分割测试特征(FAST)算法是目前应用较广泛且被认为是速度最快的特征检测算法之一^[10-11],但它不具有尺度不变性,检测到特征点后,需利用其所在局部的邻域信息进行描述。特征描述算法大致可分为基于绝对值的算法和基于比较的算法两类。其中:基于绝对值的算法主要是构造直方图,如加速鲁棒特征(SURF)^[12]描述,

虽然该方法具有尺度不变性和较高的稳健性,但计算复杂度高,效率低,满足不了现有移动设备对实时运算的需求^[13];基于比较的算法主要是构造二进制字符串,如旋转二进制鲁棒(ORB)^[14]描述子、快速视网膜膜特征(FREAK)^[15]描述子,在时间代价和计算复杂度上相较于基于绝对值的算法均大幅降低,提升了匹配速度,但此类相关算法大多以待识别图像的纹理特征提取为主,精度易受图像复杂纹理特征的影响。

针对上述两类方法存在的问题,以及目前相关领域逐渐出现的以文字为增强对象的应用需求,本文提出一种基于文字特征的识别注册方法,并根据此方法实现了以文字为识别目标的AR系统。

2 基于文字特征的识别注册方法

基于自然特征的AR识别注册包括图像预处理、特征点提取、摄像机位姿计算、虚拟物体渲染等过程。本文提出的基于文字特征的AR识别注册流程如图1所示。

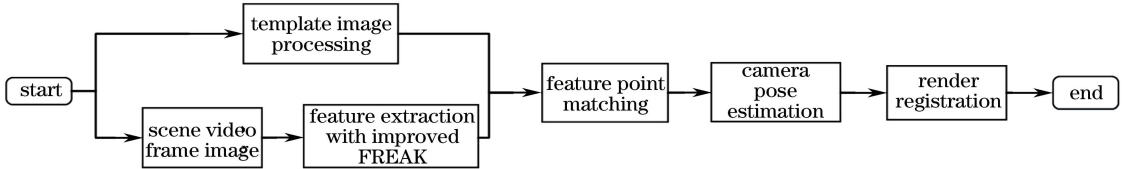


图1 基于文字特征的AR识别注册流程

Fig. 1 AR recognition registration process based on text features

基于文字特征的AR识别注册方法主要是在模板图像处理过程中对原尺度金字塔构建方法进行改进,将1.6倍降采样与2幂次采样方式相结合,以实现尺度不变性;对各尺度下的模板图像和获取到的场景视频帧图像采用基于FREAK改进的算法集中提取文字笔画特征点,并根据文字笔画角点的分布特征滤除伪特征点,以减小纹理特征点对精度的影响。完成以上步骤后,使用汉明距离将视频帧图像特征点与对应尺度下的模板图像特征点进行匹配筛选,并根据匹配点对的位置信息计算出摄像机在环境中的姿态信息,使虚拟物体可准确地渲染叠加于真实场景目标文字之上。

3 模板图像处理

特征提取过程中特征检测部分使用的基于快速特征点检测的文字定位算法(FASText)^[16]是在缺乏尺度不变性的FAST算法的基础上改进得到的,主要对文本笔画宽度半径为较小像素宽的特征点进行检测。针对一些场景下文本字体大小不同的情

况,通过对模板图像构建多尺度金字塔来保证其尺度不变性。具体步骤如下。

1) 降采样。FASText算法采用固定缩放系数 f 为1.6的降采样方法构建尺度金字塔(当系数为1.6时,文本漏检率和误检率之间相较于其他缩放系数达到了平衡)。为避免单独使用固定系数降采样或2幂次采样时出现两个尺度之间间隔太大、易出现识别失败的问题,首先对原模板图像进行一次1.6倍采样,然后对原模板图像和1.6倍采样后的图像分别不断地进行2幂次采样,并交叉放置这两类序列图。则 l 尺度下图像相对于原图像的缩放比例 S_l 为

$$S_l = \begin{cases} 2^{\frac{l}{2}}, & l \text{ is even} \\ 1.6 \times 2^{\text{Floor}(\frac{l}{2})}, & l \text{ is odd} \end{cases}, \quad (1)$$

式中: l 表示当前尺度;Floor(\cdot)表示向下取整。

2) 确定分层数。由降采样频率确定模板图像尺度数。通过计算模板图像尺度相对于摄像机视频帧图像分辨率的倍数 M ,确定对模板图像分割的最大层数 L (假设金字塔最大尺度图像分辨率刚好小于视频帧图像分辨率)。分辨率倍数 M 的计算公式为

$$M = \frac{\omega_t + h_t}{\omega_c + h_c} \times 2, \quad (2)$$

式中： ω_t 为模板图像长分辨率； h_t 为模板图像宽分辨率； ω_c 为视频帧图像长分辨率； h_c 为视频帧图像宽分辨率。

由(2)式可得对模板图像进行分割后的最大层数 L 为

$$L = \lceil 2 \times \log_2 M + 1 \rceil = \lceil 2 \times \log_2 \left(\frac{\omega_t + h_t}{\omega_c + h_c} \right) + 1 \rceil. \quad (3)$$

3) 构建金字塔。对模板图像使用确定的降采样频率进行降采样, 获得不同尺度的模板图像, 当获得最大尺度 L 后停止。

4) 提取特征点。对金字塔中不同尺度的模板图像提取特征点, 因缩放后位置发生变化, 故需将特征点位置乘以缩放倍数才能得到特征点在原始模板图像中的位置。

4 基于 FREAK 改进的文字特征提取

特征检测作为特征描述的前提, 是特征匹配

的第一阶段。FREAK 特征提取中使用的 FAST 算法主要对图像纹理特征角点进行检测。为实现对文本图像的识别注册, 使特征提取算法对文字特征点集中检测, 采用 FASText 算法替代 FAST 算法检测文本图像特征点, 并根据文字角点分布特征对伪特征点进行滤除, 以减小纹理特征对精度的影响。

4.1 FREAK 特征提取算法

FAST 特征检测主要基于分割检测判据: 以像素点为圆心, 半径为 r 的圆周上存在 n 个连续的像素点, 判断该像素点是否为角点的表达式为

$$C_{RF} = \begin{cases} 1, & \text{if } |I_p - I_k| > t \\ 0, & \text{else} \end{cases}, \quad (4)$$

式中: I_k 为圆周上任意一个像素点的灰度值 ($k=1, 2, \dots, n$); I_p 为中心像素点 p 的灰度值; t 为已知的一个很小的阈值。若 $C_{RF}=1$ 的个数多于一个给定的阈值 t' , 则将该点视为候选点, 通常选取 $t'=12$ 。FAST 特征点检测如图 2 所示。

检测到特征点后, 利用特征点所在局部邻域信息对其进行描述和表达。FREAK 描述符是模拟人眼视网膜接收图像信息而提出的方法, 如图 3 所示。

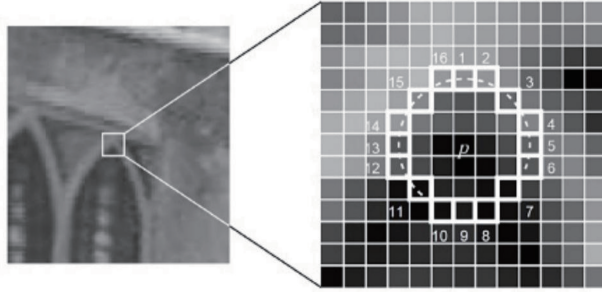


图 2 FAST 特征点检测

Fig. 2 FAST feature point detection

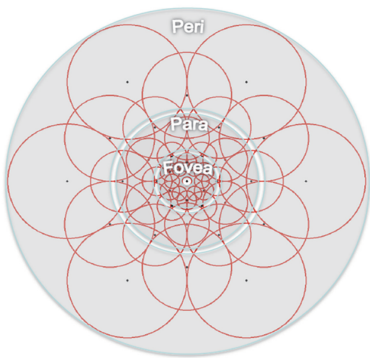


图 3 FREAK 视网膜采样模式

Fig. 3 FREAK retina sampling mode

首先对特征点邻域进行 FREAK 采样。采样点均匀分布在各圆中心, 每个圆代表一块感受野。为

了降低噪声的影响, 对每个图像进行高斯模糊处理, 圆的半径即为高斯模糊的标准差。圆与圆设置重叠部分是为了获取更多的图像信息, 使最终的描述符更独特。

FREAK 描述符是一个二进制位串, 它通过比较视网膜采样模式的图像强度来计算, 即

$$F = \sum_{0 \leq a \leq N} 2^a T(P_a), \quad (5)$$

$$T(P_a) = \begin{cases} 1, & \text{if } I(P_a^{r_1}) - I(P_a^{r_2}) > 0 \\ 0, & \text{otherwise} \end{cases}, \quad (6)$$

式中: F 表示二进制描述子; P_a 表示采样点对; N 表示期望的二进制编码; $I(P_a^{r_1})$ 与 $I(P_a^{r_2})$ 分别表示前一采样点与后一采样点的像素值。

由于特征点中的采样点较多, 因此会产生大量

匹配点对。然而,许多点对对于特征描述图像来说是无关的,因此需要对特征点对进行筛选。筛选步骤如下:

- 1) 构建矩阵 D , D 中的每一行表示一个特征点,且每一行有 C_M^2 列;
- 2) 计算矩阵 D 中每一列的平均值,若平均值接近 0.5,说明方差大,携带的信息丰富;
- 3) 根据各列中的方差由大到小进行排序;
- 4) 选取该矩阵的前 k 列作为最终的二进制描述子, k 取 512。

为了使 FREAK 算法具有旋转不变性,在检测到的每个特征点上增加方向信息。FREAK 算法通过选取 45 个距离长且对称的采样点对来计算梯度,如图 4 所示。

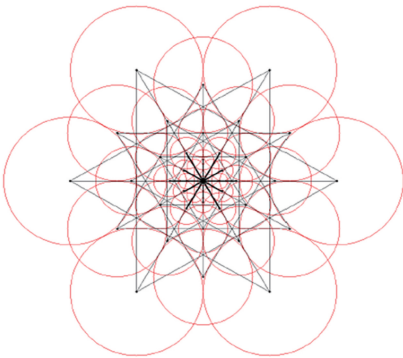


图 4 中心对称采样点对

Fig. 4 Center symmetrical sampling point pair

特征点方向的计算公式为

$$O = \frac{1}{M'} \sum_{P_0 \in G} [I(P_{\theta}^r) - I(P_{\theta}^l)] \frac{P_{\theta}^r - P_{\theta}^l}{\|P_{\theta}^r - P_{\theta}^l\|}, \quad (7)$$

式中: O 表示局部梯度信息; G 表示采样点对的集合; M' 表示采样点对的数量; P_0 表示采样点对的位置; P_{θ}^r 和 P_{θ}^l 分别为 P_0 的前一位和后一位采样点

空间坐标的二维向量,与之对应的区域的灰度均值分别为 $I(P_{\theta}^r)$ 和 $I(P_{\theta}^l)$ 。

4.2 改进的 FREAK 提取算法

为使 FREAK 特征提取算法能较准确地获取文字特征,减小纹理特征对精度的影响,在特征检测阶段对算法进行改进,将文字识别领域中的文字笔画特征点检测方法应用于此提取算法,在特征描述阶段则仍采用 FREAK 描述子对滤除后的特征点进行描述,生成二进制串。

FSAT 算法难以精确提取图像中文字弯曲的区域,且易受图像背景纹理的影响而出现大量误判角点,文字检测效果欠佳,实时性与精确性不佳,针对 FSAT 算法的这些缺陷,人们提出了以图像中文字笔画的角点作为特征点的场景文字定位算法——FASText 算法^[16]。该算法包含文字角点检测和文本方向聚合两部分,以下主要介绍本文应用到的文字角点检测部分。

文字的符号特性使得各语系文字均是基于笔画的组合,当笔画交叉或转折时即形成典型的角点。因此,FASText 在 FAST 算法角点的基础上进行改进,设计了两种新角点:SEK(笔画端点)、SBK(笔画转折点)角点,并根据特征点的中心点像素 I_p 与环形带(以两个像素宽为半径的圆)上 k ($k=12$) 个像素的差将像素点 x ($x \in \{1, 2, \dots, 12\}$) 分为 3 种:更暗的(d)、近似的(s)、更亮的(b),可表示为

$$L(p, x) = \begin{cases} d, & I_x \leq I_p - m \text{ (darker)} \\ s, & I_p - m < I_x < I_p + m \text{ (similar)}, \\ b, & I_x \geq I_p + m \text{ (brighter)} \end{cases} \quad (8)$$

式中: I_x 表示点 x 的像素; m 为判定特征点状态的域值。

根据环形带上 3 种像素的分布,定义 SEK 与 SBK。

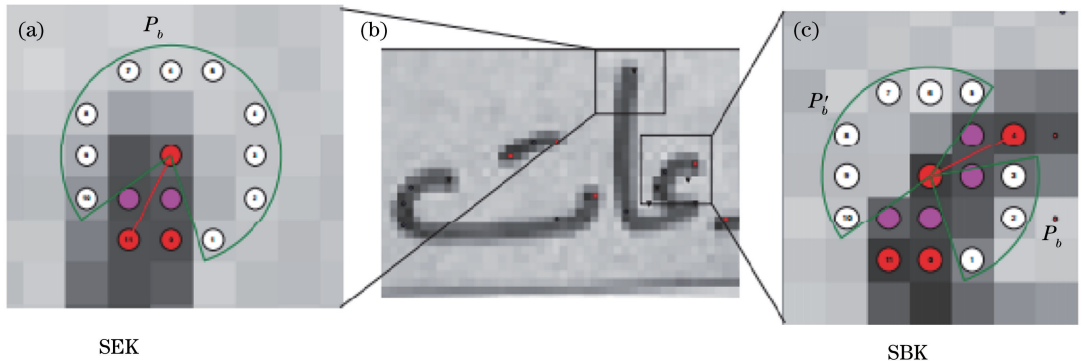


图 5 FASText 的改进角点

Fig. 5 Improved corner points of FASText

SEK 聚焦于文字笔画端点,即开始点与结束点。相较于 FAST 角点,FASText 角点不仅要求环形带上 $n(n=9)$ 个连续像素与中心点像素灰度之差大于预设阈值,而且要求剩余的 $k-n$ 个像素也接近于中心点的像素灰度值,即 $|P_s| \leq k-n$ 与 $|P_d| \geq n$ 或 $|P_b| \geq n$ 均成立,其中, P_d 、 P_s 、 P_b 为更暗、近似、更亮 3 种像素的同类集合。从图 5(a)可以看出,一般位于笔画末端的中心像素具有这种邻域分布。

SBK 聚焦于文字笔画转折点。环形带上仅有如下像素集合: P_s 、 $P_{s'}$ 、 P_d 、 $P_{d'}$ 或 P_s 、 $P_{s'}$ 、 P_b 、 $P_{b'}$, 其中 $|P_s| \leq 3$ 且 $|P_{s'}| \leq 3$, $|P_d| > 6$ 或 $|P_b| > 6$, 同时 $|P_{d'}| > 0$ 或 $|P_{b'}| > 0$ 。从图 5(b)中可以看出,在转折区中,背景像素集合 P_b 与 $P_{b'}$ 间的像素区域即为笔画转折中心点区域,而进入中心点的入口与出口像素集合分别为 P_s 与 $P_{s'}$ 。

对文本图像进行特征点检测后,尽管对图像中文字部分的检测准确度较高,但仍存在不足:摄像机获取的文本图像往往具有不同的背景,因此一些较为复杂的背景会对特征点的检测造成干扰。文本区域特征点的分布一般具有规律性,而非文本区域通常杂乱无章,无规律可寻。为了提高检测精度,可利用特征点的分布特性对伪特征点进行滤除。检测到的特征点与其环形带上相似像素点均在文本笔画上,两者间是连通的(即特征点与相似像素点之间的像素值也与特征点像素值近似),若特征点与其相似像素点之间不满足连通条件,则将该特征点过滤掉;此外,检查环

形带外围的 16 个像素点,若符合 SEK 的 P_b (或 P_d) 和 SBK 的 P_b (或 P_d) 条件,说明这个笔画点被困在里面没有延伸出去,不符合文本笔画规律,需要滤除。

将滤除后的文字特征点所在局部领域信息进行描述与表达,即使用 FREAK 描述子构造生成二进制串,完成对特征点的描述。然后,使用汉明距离对视频帧图像与模板图像特征点进行匹配,采用渐进一致采样(PROSAC)^[17]方法剔除误匹配点对,提升效率。进而计算摄像机位姿,在摄像机视角下的渲染坐标系中完成注册,使虚拟物体正确放置于真实场景中,实现虚实融合。

5 实验结果及分析

实验环境为 Windows 7 系统, Intel i7-6700 CPU, 32 GB RAM, 采用 VS2015、OpenCV3.3.0、OpenGL 等软件平台。以 AR 在教育领域中的应用为例,选取 5 组图像(每组包含了 3 幅具有相同文字、不同复杂度纹理的文本图像)进行特征匹配对比实验,并选取其中的一组对比图进行分析说明。

对各组图像进行尺度和旋转变化处理后,分别采用 SURF、ORB、FREAK 和改进的 FREAK 算法进行特征匹配。表 1 是 4 种算法对 3 类文本图像特征匹配后平均运行时间、平均特征点对数及特征匹配精度的数据统计,此外对改进的 FREAK 算法较其他三种算法的平均运行时间及特征点对数的减少程度进行了计算比较。

表 1 各算法平均运行时间、平均特征点对数和匹配精度的比较

Table 1 Comparison of average running time, average number of feature point pairs, and matching precision of algorithms

Parameter	SURF			ORB			FREAK			Improved FREAK		
	A	B	C	A	B	C	A	B	C	A	B	C
Average running time /ms	36.6	55.2	57.6	34.5	38.2	39.3	33.8	36.5	37.6	31.8	32.7	33.1
Contrast of average running time /%	-13.1	-40.7	-42.5	-7.8	-14.3	-15.7	-5.9	-10.4	-11.9	-	-	-
Average number of feature point pairs	358	572	607	336	509	535	324	469	498	302	338	342
Contrast of average number of feature point pairs /%	-15.6	-40.9	-43.6	-10.1	-33.5	-36.0	-6.7	-27.9	-31.3	-	-	-
Precision /%	69.6	64.2	62.3	76.9	75.5	74.6	78.7	76.2	75.4	87.8	85.4	84.6

表 1 中根据图像纹理的复杂度将 5 组图像分为 A、B、C3 类,其中, A 类为单一文本, C 类较 B 类图像的纹理更复杂。从表 1 可以看出: SURF 算法对 3 类图像的匹配运行时间最长, 特征点对数最多; ORB 算法与 FREAK 算法均因包含了部分匹配错误的特征点对而缩短了运行时间。因样本较少,除了 SURF 算法识别精度较低外,其他算法的识别精

度均保持在 74% 以上。综合比较后可知,在对文本图像内容识别时,所提算法在耗时、特征点对数量和精度方面均优于其他 3 种算法。下面以其中一组文本图像的特征匹配结果图为例进行说明。

从不同算法对只具有文字的图像的特征匹配结果(图 6)可以看出,所提算法对特征点的检测更集中于文字笔画的端点和转折点,获取匹配得到的特征

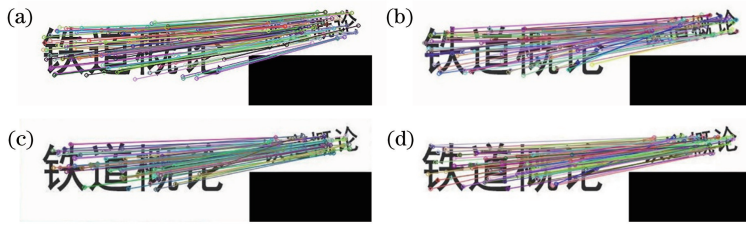


图 6 A 类文本图像的匹配结果。(a) SURF 算法;(b) ORB 算法;(c) FREAK 算法;(d)改进的 FREAK 算法
Fig. 6 Matching results of text images of class A. (a) SURF algorithm; (b) ORB algorithm; (c) FREAK algorithm; (d) improved FREAK algorithm

点对相比其他 3 种算法更准确,且数量最少。

B 类文本图像相较于 A 类文本图像具有较复杂的纹理,对比结果(图 7)显示,所提算法提取匹配到的特征点对主要集中于“铁道概论”文字上,而其他 3 种算法提取匹配到的特征点则包含了大量的纹理特征,文字特征较少。

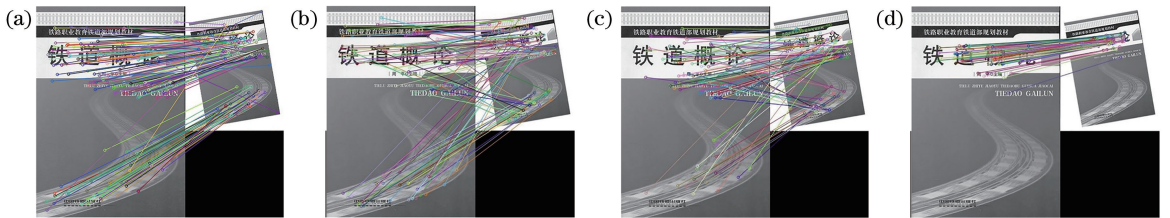


图 7 B 类文本图像的匹配结果。(a) SURF 算法;(b) ORB 算法;(c) FREAK 算法;(d)改进的 FREAK 算法
Fig. 7 Matching results of text images of class B. (a) SURF algorithm; (b) ORB algorithm; (c) FREAK algorithm; (d) improved FREAK algorithm

C 类文本图像在 3 类文本图像中的纹理最复杂,结果(图 8)同样显示所提算法主要集中于对“铁道概论”文字特征的提取。但从此对比图中也可看出,随着纹理的复杂化,所提算法对文字笔画特征角点的提取效果逐渐变差,这有待于进一步改进。

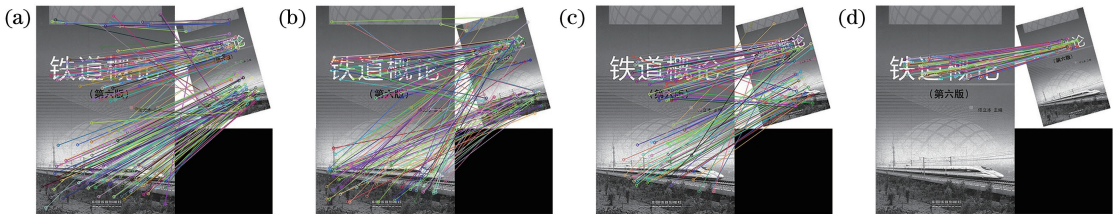


图 8 C 类文本图像的匹配结果。(a) SURF 算法;(b) ORB 算法;(c) FREAK 算法;(d)改进的 FREAK 算法
Fig. 8 Matching results of text images of class C. (a) SURF algorithm; (b) ORB algorithm; (c) FREAK algorithm; (d) improved FREAK algorithm

最后,实现了基于文字特征的 AR 系统,并以“铁道概论”3 幅文本图像作为增强对象进行识别注册。图 9~11 为所提方法分别在不同视角、不同方

向距离,以及不同的图像被部分遮挡情况下的识别注册结果。表 2 是选取的 3 幅图像在 4 种算法集成的系统下实现每个步骤的平均耗时。

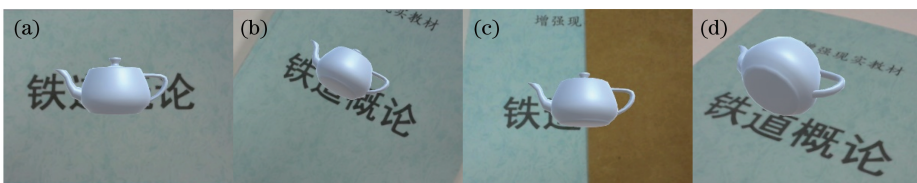


图 9 图像 A 的识别注册。(a)正面注册;(b)距离、旋转变注册;(c)遮挡注册;(d)视角变化注册
Fig. 9 Recognition registration for image A. (a) Positive registration; (b) distance and rotation change registration; (c) occlusion registration; (d) perspective change registration



图 10 图像 B 的识别注册。(a)正面注册;(b)距离、旋转变化注册;(c)遮挡注册;(d)视角变化注册

Fig. 10 Recognition registration for image B. (a) Positive registration; (b) distance and rotation change registration; (c) occlusion registration; (d) perspective change registration

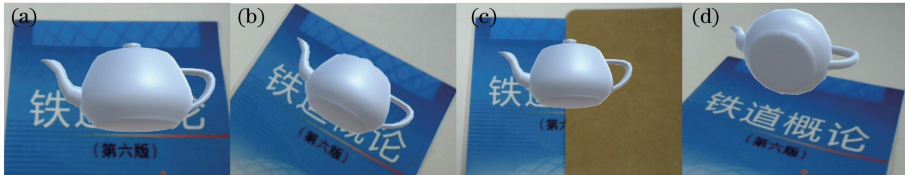


图 11 图像 C 的识别注册。(a)正面注册;(b)距离、旋转变化注册;(c)遮挡注册;(d)视角变化注册

Fig. 11 Recognition registration for image C. (a) Positive registration; (b) distance and rotation change registration; (c) occlusion registration; (d) perspective change registration

表 2 不同算法的平均处理时间

Table 2 Average processing time of different algorithms

ms

Algorithm	Image processing	Feature extraction	Feature matching	Pose estimation	Model rendering	Total time
SURF	1.3	21.6	18.1	8.7	6.4	56.1
ORB	1.3	18.2	17.5	9.8	6.4	53.2
FREAK	1.3	16.8	15.4	9.8	6.4	49.7
Improved FREAK	1.3	15.1	13.8	9.8	6.4	46.4

从 3 幅文本图像的注册效果可以看出,所提方法在几种变换操作下均可以完成对文本图像的准确注册,将虚拟物体叠加于图像文字之上。分析对比相关数据可知,随着图像纹理的复杂化,各算法的平均处理时间均随之增加,其中,所提算法的耗时较少,每帧图像的总耗时为 46.4 ms,处理速度约为 21 frame/s,虽不能很好地满足实时性要求,但更适用于对文本图像的识别注册,可满足部分领域对文本内容增强的应用需求。

6 结 论

提出了一种基于文字特征的识别注册方法,并实现了针对文字进行识别注册的 AR 系统。通过对比实验可看出,在针对文本图像的识别中,所提算法在时间与精度上均优于其他 3 种对比算法,在待识别文本图像发生变化时仍能较准确地识别注册,表现出了较好的增强效果。

参 考 文 献

[1] Bimber O, Raskar R. Spatial augmented reality: merging real and virtual worlds[M]. Boca Raton:

CRC Press, 2005.

[2] Zhou Z, Zhou Y, Xiao J J. Survey on augmented virtual environment and augmented reality [J]. Scientia Sinica Informationis, 2015, 45(2): 157-180. 周忠,周颐,肖江剑. 虚拟现实增强技术综述[J]. 中国科学:信息科学, 2015, 45(2): 157-180.

[3] Yu D G, Jin J S, Luo S H, et al. A useful visualization technique: a literature review for augmented reality and its application, limitation & future direction[M] // Huang M, Nguyen Q, Zhang K. Visual information communication. Boston, MA: Springer, 2009: 311-337.

[4] Gui Z W. Study on scene identification and registration technology for mobile augmented reality [D]. Beijing: Beijing Institute of Technology, 2014. 桂振文. 面向移动增强现实的场景识别与跟踪注册技术研究[D]. 北京: 北京理工大学, 2014.

[5] An Z, Xu X P, Yang J H, et al. Three-dimensional tracking registration method based on semantic object matching[J]. Acta Optica Sinica, 2018, 38(12): 1212002. 安喆, 徐熙平, 杨进华, 等. 基于语义目标匹配的 3 维跟踪注册方法 [J]. 光学学报, 2018, 38(12):

- 1212002.
- [6] Zhou F, Duh H B L, Billingham M. Trends in augmented reality tracking, interaction and display: a review of ten years of ISMAR[C] // 2008 7th IEEE/ACM International Symposium on Mixed and Augmented Reality, September 15-18, 2008, Cambridge, UK. New York: IEEE, 2008: 193-202.
- [7] Kim K, Park N Y, Woo W. Vision-based all-in-one solution for augmented reality and its storytelling applications[J]. *The Visual Computer*, 2014, 30(4): 417-429.
- [8] Gui Z W, Liu Y, Chen J, et al. Online learning of tracking and registration based on natural scenes[J]. *Journal of Software*, 2016, 27(11): 2929-2945.
桂振文, 刘越, 陈靖, 等. 基于自然场景在线学习的跟踪注册技术[J]. *软件学报*, 2016, 27(11): 2929-2945.
- [9] An Z, Xu X P, Yang J H, et al. Design of augmented reality head-up display system based on image semantic segmentation[J]. *Acta Optica Sinica*, 2018, 38(7): 0710004.
安喆, 徐熙平, 杨进华, 等. 结合图像语义分割的增强现实型平视显示系统设计与研究[J]. *光学学报*, 2018, 38(7): 0710004.
- [10] Rosten E, Drummond T. Machine learning for high-speed corner detection[M] // Leonardis A, Bischof H, Pinz A. *Computer vision - ECCV 2006. Lecture notes in computer science*. Berlin: Springer, 2006, 3951: 430-443.
- [11] Rosten E, Porter R, Drummond T. Faster and better: a machine learning approach to corner detection[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, 32(1): 105-119.
- [12] Bay H, Tuytelaars T, van Gool L. Surf: speeded up robust features[C] // Leonardis A, Bischof H, Pinz A. *Computer vision - ECCV 2006. Lecture notes in computer science*. Berlin: Springer, 2006, 3951: 404-417.
- [13] Xiong J Y, Dai M. Design of tracker for mobile smart devices[J]. *Optics and Precision Engineering*, 2017, 25(12): 3152-3159.
熊晶莹, 戴明. 适应移动智能设备的目标跟踪器[J]. *光学精密工程*, 2017, 25(12): 3152-3159.
- [14] Rublee E, Rabaud V, Konolige K, et al. ORB: an efficient alternative to SIFT or SURF [C] // 2011 International Conference on Computer Vision, November 6-13, 2011, Barcelona, Spain. New York: IEEE, 2011: 2564-2571.
- [15] Alahi A, Ortiz R, Vanderghyest P. FREAK: fast retina keypoint [C] // 2012 IEEE Conference on Computer Vision and Pattern Recognition, June 16-21, 2012, Providence, RI, USA. New York: IEEE, 2012: 510-517.
- [16] Buta M, Neumann L, Matas J. FASText: efficient unconstrained scene text detector [C] // 2015 IEEE International Conference on Computer Vision (ICCV), December 7-13, 2015, Santiago, Chile. New York: IEEE, 2015: 1206-1214.
- [17] Chum O, Matas J. Matching with PROSAC: progressive sample consensus [C] // 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), June 20-25, 2005, San Diego, CA, USA. New York: IEEE, 2005: 8588877.