

基于对焦清晰度的双分辨率相机变焦算法研究

金哲彦, 徐之海, 冯华君, 李 奇

(浙江大学 现代光学仪器国家重点实验室, 浙江 杭州 310007)

摘要: 双分辨率相机同时兼顾大视场和高分辨, 利用同轴光学固定结构避免了变焦镜头运动部件带来的诸多问题, 在深空探测目标跟踪和手机等智能终端上具有应用价值。针对现有的基于深度学习的双分辨率图像变焦算法的速度慢, 信息量没有提升, 图像网络结构适配性差和图像信息修复的伪造性等问题尝试性地提出了加入基于深度信息的解决办法。论证将图像对焦清晰度作为深度信息引入双分辨率变焦算法的可行性, 探讨对焦深度信息检测精度与效果, 分别测试深度学习和基于深度信息的传统方法的双分辨率变焦算法, 得到了一种不影响正常成像速度, 内存开销降低 35% 和算法复杂度降低 60%, 超分辨率信息真实可靠, 图像结果评价提升 10% 到 50% 的全新算法。

关键词: 双分辨率成像; 光学变焦; 图像融合; 图像清晰度评价; 光学超分辨率
中图分类号: TP751 **文献标志码:** A **DOI:** 10.3788/IRLA20190463

Research on zoom algorithm based on dual resolution camera

Jin Zheyang, Xu Zhihai, Feng Huajun, Li Qi

(State Key Laboratory of Modern Optical Instrumentation, Zhejiang University, Hangzhou 310007, China)

Abstract: The dual resolution camera combines large field of view and high resolution. The coaxial optical fixed structure avoids many problems caused by the moving parts of the zoom lens, and has potential application value in deep space detection target tracking and smart terminals such as mobile phones. The existing dual resolution image zoom algorithm based on deep learning has slow speed, no increase in information, poor adaptability of image network structure and forgery of image information restoration. To join a solution in normal image algorithm based on depth information was tentatively proposed in this paper. The feasibility of using image focus sharpness as depth information into the dual resolution zoom algorithm was demonstrated, the accuracy and effect of focus depth information detection was explored by Laplace image evaluation, and the dual resolution zoom algorithm of deep learning and traditional methods based on depth information was tested. The new excellent zoom algorithm used the space during normal focus of the camera which did not affect the imaging speed of normal images. The algorithm complexity is reduced by 60%, the computational memory overhead is reduced by 35%. The spatial relationship super-resolution information is more realistic and reliable, and the image result evaluation is improved by 10% to 50%.

Key words: dual resolution imaging; optical zoom; image fusion; image sharpness evaluation; optical super-resolution

收稿日期:2020-01-13; 修订日期:2020-02-15

基金项目:十三五民用航天预先研究项目(D040301)

作者简介:金哲彦(1996-),男,博士生,主要从事成像与图像处理方面的研究。Email:11930051@zju.edu.cn

导师简介:李奇(1973-),男,副教授,博士,主要从事成像与图像处理方面的研究。Email:liqi@zju.edu.cn

0 引言

光学超分辨成像是光学研究的热门领域,无限逼近甚至超过光学成像的理论分辨率成为了整个光学领域最重要的课题之一。双分辨相机的本质问题其实是光学成像超分辨的应用。航天双分辨相机利用在镜头组里的大棱镜组分光实现同轴不同尺寸和分辨率的双像面成像,可以减小航天器光学系统的载荷并提高光学系统的稳定性,做到“大视场追踪小视场看清”。浙江大学光学成像课题组已经将航天双分辨相机应用于嫦娥四号中继星并取得了不俗的成绩。双分辨相机的数字变焦技术,主要是提高双分辨相机的工作性能,兼顾大视场和高精度。同时对于智能手机等个人消费品而言,两个以上不同焦距的定焦摄像头也可以组成近似同轴的双分辨相机系统。优秀的双分辨变焦算法可以大大提高智能手机的拍摄水平,通过手机自身强大的运算能力和多定焦镜头来弥补自身为了便携性所牺牲掉的光学变焦特性。

大部分双分辨率变焦算法还停留研发阶段。现有算法主要分三种,基于卷积神经网络(CNN)特征提取的双焦相机连续数字变焦^[1],基于光流深度学习的双分辨率算法,基于高频信息重构的纹理匹配和小波变换算法^[2]。在时域和频域两个方面分别利用传统方法和深度学习的方法或者多通道信息融合的办法将窄视场高分辨图像的信息提取出来,提高宽视场低分辨率的图像质量。同时国内外研究者提出了很多优秀的 RefSR^[3](有参考的超分辨)算法:2018 年的基于 CrossNet 的 RefSR^[4],2019 年 3 月的 SRNTT 算法等^[5]。

这些算法的使用在双分辨相机当中取得了较好效果(原因是参考图像是同轴的),但也出现了计算速度偏慢,虚假信息参考等问题,这也是文中算法要解决的问题。文中将引入对焦过程中的清晰度变化作为双分辨算法的约束和辅助计算条件。

1 算法理论原理

1.1 双分辨相机原理

无论激光对焦、相位对焦、反差对焦,相机必定在对焦过程中提取信息。在正常对焦过程当中加入深度信息的提取对于成像速度没有新增的负担。通过景深公式(1)可以看出当相机镜头正常的对焦之后,其他参数不变,物距 d 越大,景深 Z 越大,前景深 Z 比后景深 Z 要小^[6]。因此在文中算法中,对焦点前的深度信息会被适当压缩。焦点后的深度信息会较为松散,在物距较长的拍摄状态下景深的宽容度也较大,深度信息矩阵也可以在更大的动态范围内工作。小物距高精度,大物距大量程恰恰非常适合文中算法。

$$Z = \frac{4rf^2dD}{4r^2(d-f)^2 - f^2D} \quad (1)$$

1.2 对焦过程中的视场变化问题

传感器尺寸不变,相机在进行对焦过程中由于物距的变化视场角会有略微变化。如公式(2)中焦距 f 不变,视场角 θ 随着物距 d 正相关变化。图 1 中显示了对焦过程中的视场区域的变化,可以看到图像边缘有一些裁切。



图 1 对焦过程中的视场变化

Fig.1 Field of view change during focusing

$$2\theta = 2\arctan\left(\frac{d}{2f}\right) \quad (2)$$

因此未经处理的图像深度矩阵并不能直接使用,解决这类问题的办法有两种。基于算法的尺度不变特征变换 (SIFT) 或加速稳健特征 (SURF) 配准和利用相机马达硬件信息^[7]。SIFT 算法的匹配准确率高于 SURF 算法, SURF 算法的匹配速度高于 SIFT 算法。两者算法对于图像的配准效果很好,但是文中的方法只记录了简单的深度信息矩阵,相比于完整的图像信息量较少,可能对特征点配准有不利影响。而对于完全同轴的双分辨率图像只需要根据相机的镜头模组的音圈马达变焦信息进行合适的裁切即可,从而大大降低了算法复杂度。另外需要注意的是不同的对焦图像存在不同的像差,但定焦镜头和同轴因素使像差在纹理信息匹配中影响不大。手机等设备在对焦过程中的抖动可以参考最新的全景稳像算法来解决^[8]。

1.3 图像清晰度评价函数的选用

主流图像清晰度评价函数主要是用于评价图像质量和相机对焦效果的。图像清晰度评价函数可分为梯度函数和频谱函数^[9],部分学说认为熵函数也有应用价值也可分为一类。常用的梯度函数有 Tenengrad 函数、能量梯度函数、Brenner 函数、方差函数。相同情况下,频谱运算时间长,但精度高。梯度函数运算是速度快,但精度低。一个好的评价函数需要具有单峰性、无偏性、灵敏性。文中测试发现在对于不同对焦情况下,采用 Laplace 算子、能量梯度算子和 Brenner 梯度法较好,而方差法效果较差, Laplace 算子是最适合文中深度信息提取的图像清晰度评价函数。但要注意的是,虽然 Laplace 评价函数的精度

高但是抗噪声性能较差,对于噪点多的图像建议先去噪或者使用别的评价函数。

1.4 对焦清晰度和图像纹理关系

1.4.1 常规成像关系

对大部分图像来说都会有前景和背景,对于同一深度的物体人们可以简单地认为是同一个纹理区域带。不同的物体有不同纹理,不同的物体一般也有不同的空间深度关系。利用相机成像过程中本来就有的对焦信息进行提取,使用对焦清晰度便可以将不同纹理区域提取出来,并计算不同的纹理之间的关系。

图 2 中对某一常规景物进行拍摄的实例,该图是实验演示图像组,真实拍摄的时候相机不需要这部分图像信息,只需要在对焦过程当中提取图 2(b) 部分灰度内容即可。

可以发现图 2 经过后转成灰度图像,高像素值代表对焦准确区域。在某一景深的位置范围内,不同图像当中的不同像素值也代表了不同的深度信息。将这些信息结合便可以得到这个拍摄场景不同物体的深度关系。通过图 2(b) 可以看到深度逐渐从右侧的树丛向左移动,最后移动到窗框和鞋子,甚至对于复杂树丛也可以分析出层叠纹理的远近位置关系。

1.4.2 纹理边缘对深度提取的影响

任何一种图像清晰度评价函数应用在文中的算法内都有缺陷。基于纹理变化的图像清晰度函数只有明显的纹理变化区域才会体现较高的图像清晰度评价函数值,对于普通的纯色小块区域图像清晰度评价函数值较低,例如由远到近的墙面纹理信息。单纯的图像清晰度矩阵和文中想要的深度信息存在差距。



(a)

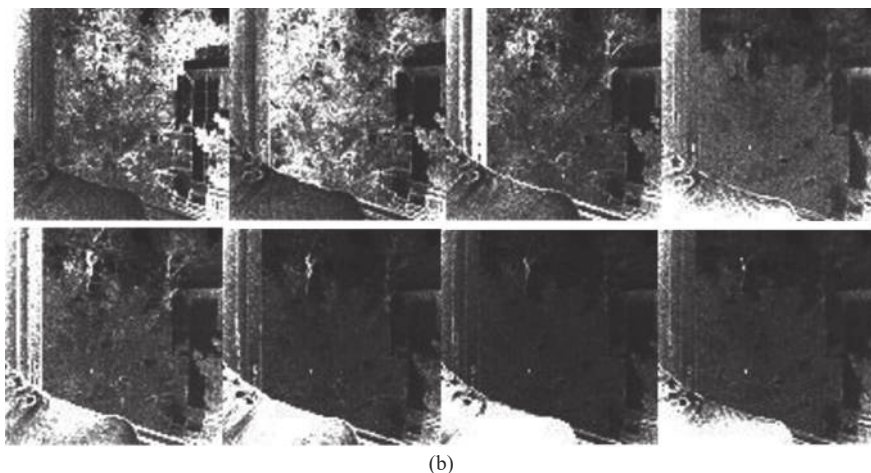


图 2 基于对焦实拍图像的深度信息提取

Fig.2 Depth information extraction based on focus real shot image

然而这种图像处理结果上的差距恰恰是文中方法的优势。Laplace 算子不仅是图像清晰度评价还是边缘信息提取函数。双分辨率变焦算法本质上是将高频纹理信息进行迁移。大量的纯色无纹理区域的无法定位对于双分辨率变焦算法没有影响,因为无纹理区域进行纹理迁移或图像小块匹配没有意义。使用基于对焦信息的图像清晰度函数不仅仅是为了找到不同图像中物体的大概位置关系,更是为了找到不同纹理的大概位置关系。这也决定了文中方法的速度优势和信息真实性。

1.4.3 纹理信息对深度提取的影响

图像清晰度函数是基于图像的纹理关系的图像评价函数。文中通过对焦的清晰度来判断物体的实际位置。但面对焦外多纹理区域和焦内低纹理区域,利用 Laplace 函数评判效果较差。容易把焦外多纹理区域判断成为焦内的同一个物体。图 3(a) 为大景深原场景图,图 (b)、(c)、(d) 为小景深下对焦在远中近的深度信息图像。图 3(a) 中近处衣物即便在最远对焦情况下依然亮度最高的。单纯一张图像无法区分没有对焦上的密纹理衣物和对焦上的光滑物体。

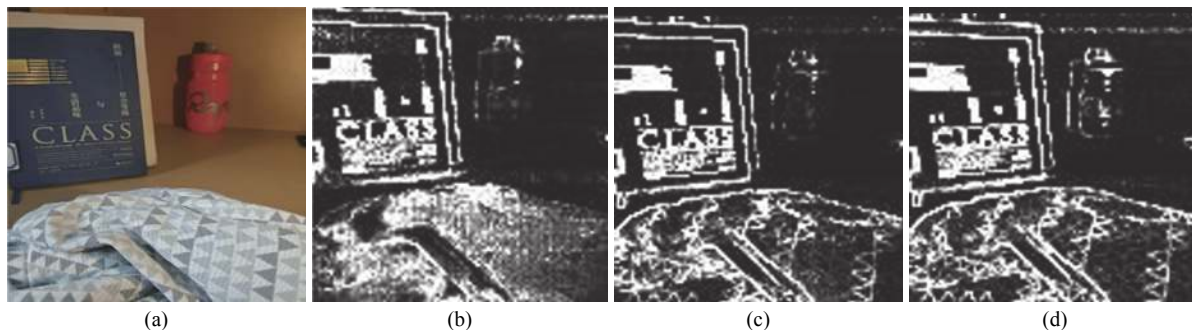


图 3 对焦位置不同对深度提取的影响

Fig.3 Different focus positions impact depth information extraction

基于上述纹理对于深度信息提取的影响,文中提出了差分 Laplace 的处理方法及基于前后两个不同的对焦点的深度信息矩阵。前后两个矩阵的匹配精度非常关键,文中 1.2 节提出的基于相机马达的同轴匹配就是应用在这里,其精度决定了差分后的检测效果。

1.5 算法流程

1.5.1 主算法流程

文中图像算法流程如图 4 所示,普通的双分辨率图像小块匹配算法只有右边的一列,将高分辨窄视场和低分辨率宽视场的图像根据倍率关系 paste 后,选择合适的图像小块尺寸,根据合适的图像小块匹配标

准,把已有的图像小块和基于深度学习的动态图像小块合成。文中提出的工作在流程图的左侧,主要是通过配准过的基于差分的 Laplace 深度矩阵来寻找图像中的深度信息,作为图像匹配距离的控制依据和图像

小块生成依据,从而大大提高图像的信息量和匹配效率,拒绝无用和错误匹配。文中算法需要相机正常对焦过程当中的对焦信息,不同于普通的多帧成像,这里只需要不同对焦段的对焦信息。

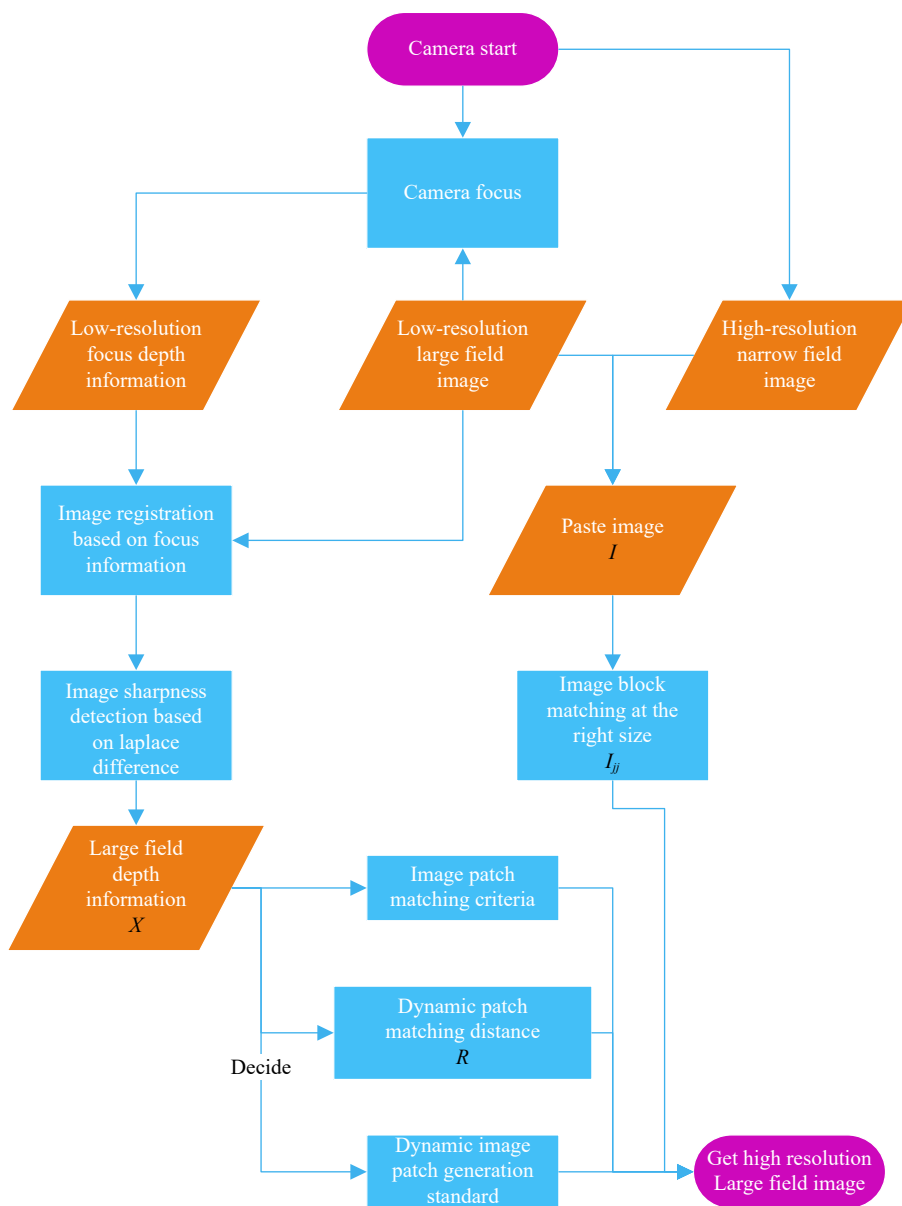


图 4 主算法流程图

Fig.4 Main algorithm flow chart

图像小块的匹配标准有很多,越简单的图像小块匹配算法,越能体现文中监督方法的有效性。文中采用颜色三通道 phash 小块匹配方法,该方法能同时解决颜色和纹理匹配两大问题。小块的拉伸,旋转都根据具体图像具体设定。当然也有很多不错的图像匹配算

法,例如离散余弦变换 DCT 匹配, Gram 纹理尺寸匹配, BM3D 加速的图像小块匹配等,都可以应用在文中的算法当中。

1.5.2 基于深度信息的图像小块预测流程

以往图像小块的匹配算法主要分两种:直接替换

法和基于感受野的深度图像恢复方法。直接替换法是利用高清图理的图像小块去替换低纹理区域类似的图像小块,这样画面更加锐利但信息量没有提升,反而块与块之间的差异更小了。基于感受野的深度图像恢复方法可以认为是在高神经网络层的直接替换法,在将图像恢复到底层,来提高画面质量。看似提高了信息量但是并没有基于传感器的实际信息。

因此,文中提出了基于对焦清晰度的梯度预测的图像小块动态预测与生成。主要的做法是将在目标块的图像动态范围之内寻找满足图像纹理匹配标准的小块,并把这些新小块的深度信息挖掘出来,通过加权平均等方式调整纹理和颜色来生成新的图像小块。本质上是把基于对焦的深度信息放入到小块纹理关系当中。具体基于深度信息的图像小块算法流程算法如下:

- (1) 从深度矩阵 X 中计算图像的 R (R 为深度矩阵 X 相近区域的范围)。
- (2) 从图像 I 中找需要匹配的图像小块 I_{ij} (I_{ij} , X_{ij} 为 ij 位置的图像小块和深度矩阵的值)。
- (3) 从深度矩阵 X 当中找到 I_{ij} 对应的 X_{ij} 。
- (4) 寻找 X_{ij} R 范围内的所有 X_{ab} 。
- (5) 在中部高分辨率图像部分找到符合 $|X_{ab}-X_{ij}| < \min|X_{ab}-X_{ij}| * R$ 的 I_{ab} 。
- (6) 检验 I_{ab} 和 I_{ij} 是否符合基础匹配条件。
- (7) 根据 $|X_{ab}-X_{ij}|$ 作为权重匹配所有 I_{ab} 。
- (8) 用第 (7) 步的结果替换 I_{ij} 。
- (9) 返回第 (2) 步,第 (3)~(6) 步无法执行返回第 (2) 步。

1.6 实验流程

算法: 基础的图像处理过程

- (1) 拍摄图像 4 032×3 016(2 张 72dpi 图像)。
- (2) resize 到 4 000×3 000 或 500×500。
- (3) crop(0.25 0.25 0.5 0.5)。
- (4) bicubic imresize 第二步一半分辨率。
- (5) 同轴 paste 第 (3), (4) 步结果得到图像 I 。
- (6) 对于第 (4) 步的结果计算图像深度矩阵。

成像设备: 智能手机, 主摄像头: 光圈 $f/1.8$, 4 mm 焦距, 图像分辨率为 4 032×3 016。副摄像头: 光圈 $f/2.6$, 6 mm 焦距, 图像分辨率为 4 208×3 120。无论主

摄像头、副摄像头生成的图像均为 72dpi 24 位深度 sRGB 色域空间的图像。

计算设备: AMD Ryzen 5 1600+B450M+16 G(DDR4 2666)+GTX 1060 5 GB+L8H 256 GB SSD(代码放置)+Samsung 850EVO SSD(图片放置)。

软件环境: Python 3.6(IDE) 和 Matlab。

图像采集: 实验基础的图像采集过程如下文算法。高分辨率小视场和低分辨率大视场的图像尺寸为 1:2。第 (3) 步的图像也可以使用手机副摄像头的成像结果经过融合来代替。

2 结果与讨论

2.1 图像小块的尺寸选择

选择合适的图像小块尺寸是文中方法成功的关键。图像小块的尺寸选择分两个部分,深度信息的小块尺寸和图像匹配小块尺寸的大小。

为了避免无效的重复计算和边缘信息丢失,一般匹配图像小块尺寸小于深度图像尺寸、小于原图像尺寸,取整数倍的比例缩放。因为深度信息矩阵的图像小块只需要找到大概的轮廓和深度为后续的匹配提供信息即可,精度要求远低于图像匹配小块。根据表 1 可以发现深度矩阵相对复杂度略低于矩阵尺寸平方增长。文中方法下相机应分别获取未对焦和对焦后的两张图像。在对焦过程中对前一张未对焦获得图像计算深度矩阵。应根据相机对焦时间和算力确定最大的深度信息尺寸。如图 5 所示,图像在 50~100 尺寸已经可以起到区分深度的作用了,100~200 可以为图像小块动态匹配提供依据,低于 50 或高于 500 不建议使用文中的方法。文中方法是用光学方法提高计算速度,不应本末倒置,降低正常成像速度。

表 1 不同尺寸深度矩阵的运算时间(单位: s)

Tab.1 Calculation time of different size depth matrix (Unit:s)

Size depth matrix size	4 000×3 000 (original image)				
	20×20	50×50	100×100	200×200	500×500
Computing time	6.61	16.11	54.60	188.84	1 114.69

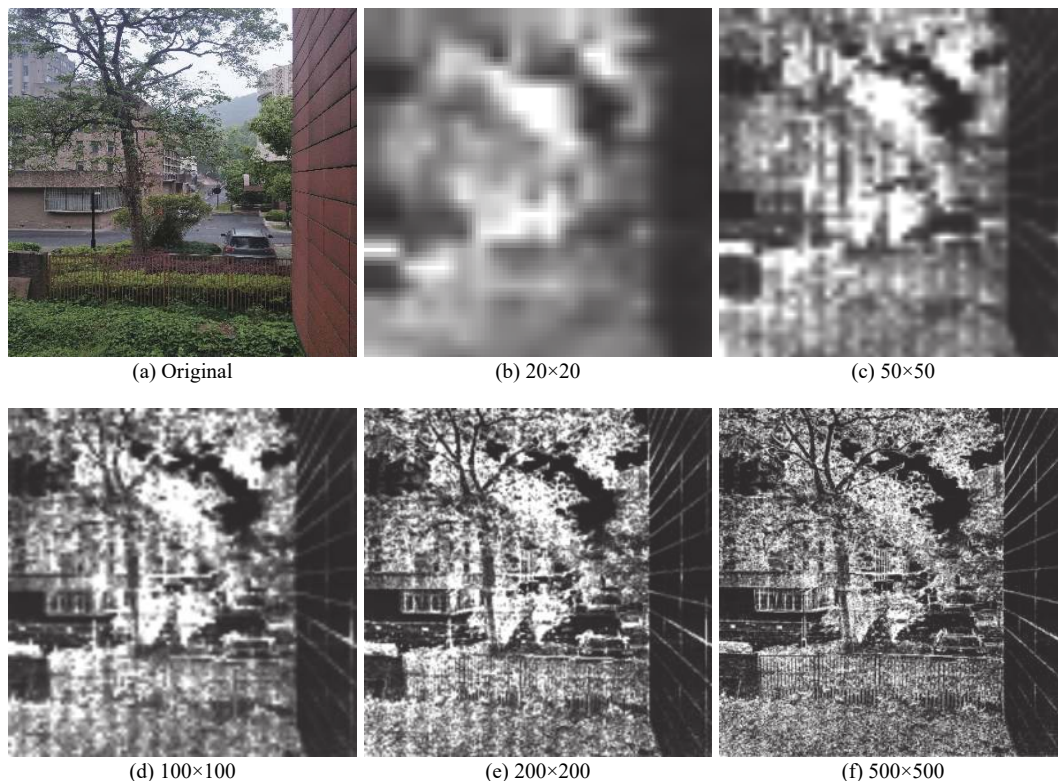


图 5 不同尺寸的深度图像小块比较

Fig.5 Comparison of different size depth image patches

2.2 基于对焦清晰度的深度检测

深度提取的性能是文中提出的变焦算法的核心。越多的深度信息图像对于区分物体的效果更好，越多的深度信息矩阵对于去除文中 1.4.3 节中图像纹理的副作用越好。文中的方法最低要求一个对焦过程的深度矩阵和正常双分辨图像，这样才能通过差分来去除纹理的副作用，并且计算出前后的位置关系。并且两张矩阵的对焦点差距越大效果越好。但文中建议在相机正常对焦的过程当中应该以正常对焦为高优先级，因为深度信息矩阵的精度要求较低，且引入深度矩阵是为了解决成像计算速度慢的问题，不能背道而驰。

通过前面图 3 的对比，不难发现图像清晰度函数不仅可以勾勒纹理区域还可以区分成像物体的远近。前景的布料在对焦近处时清晰度函数的值很高，在对焦中等深度时次之，对焦远处时最弱。中景物体书本在对焦中等深度时最为锐利，前后景的图像清晰矩阵差距不大。后景的水壶由于不像前面两种景物纹理丰富，图像清晰度值相对较低，但也能看的出在对焦近处时远没有对焦远处时深度矩阵锐

利。基于对焦清晰度的图像深度矩阵完全检测出了不同物体，并且为双分辨相机变焦提供了所有需要的信息。

2.3 基于对焦清晰度的双分辨算法图像结果

文中的算法处理结果如图 6 所示。通过图像纹理细节，可以明显看到图像质量的提升。图 6(a) 为双分辨率 paste 图像，图 6(b)(c)(d) 分别为近中远处图像处理结果，图 6(e) 为 SRCNN 结果，图 6(f) 为真实 (label) 图像。由表 2 可知，文中方法的 PSNR 不如 SRCNN (PSNR 和人眼实际感受上差异已经是图像超分辨领域的共识)，但在图像的结构相似性 SSIM 上是完胜 SRCNN 的，可见 SRCNN 在 PSNR 上的领先是付出巨大代价的。而文中的方法 Laplace 算子评价价值相比于原图有较大提升。

图 6 显示了 500×500 的图像大尺寸图像小块的匹配结果，4 000×3 000 尺寸显示效果不明显。经过测试 4 000×3 000 的 paste 图像应用文中的方法，在小尺寸图像小块的情况下 (相当提高了图像小块匹配精度)，图像处理结果更加细腻。平均图像 Laplace 从 232.25 提升到 386.28，平均提升超过 60%。

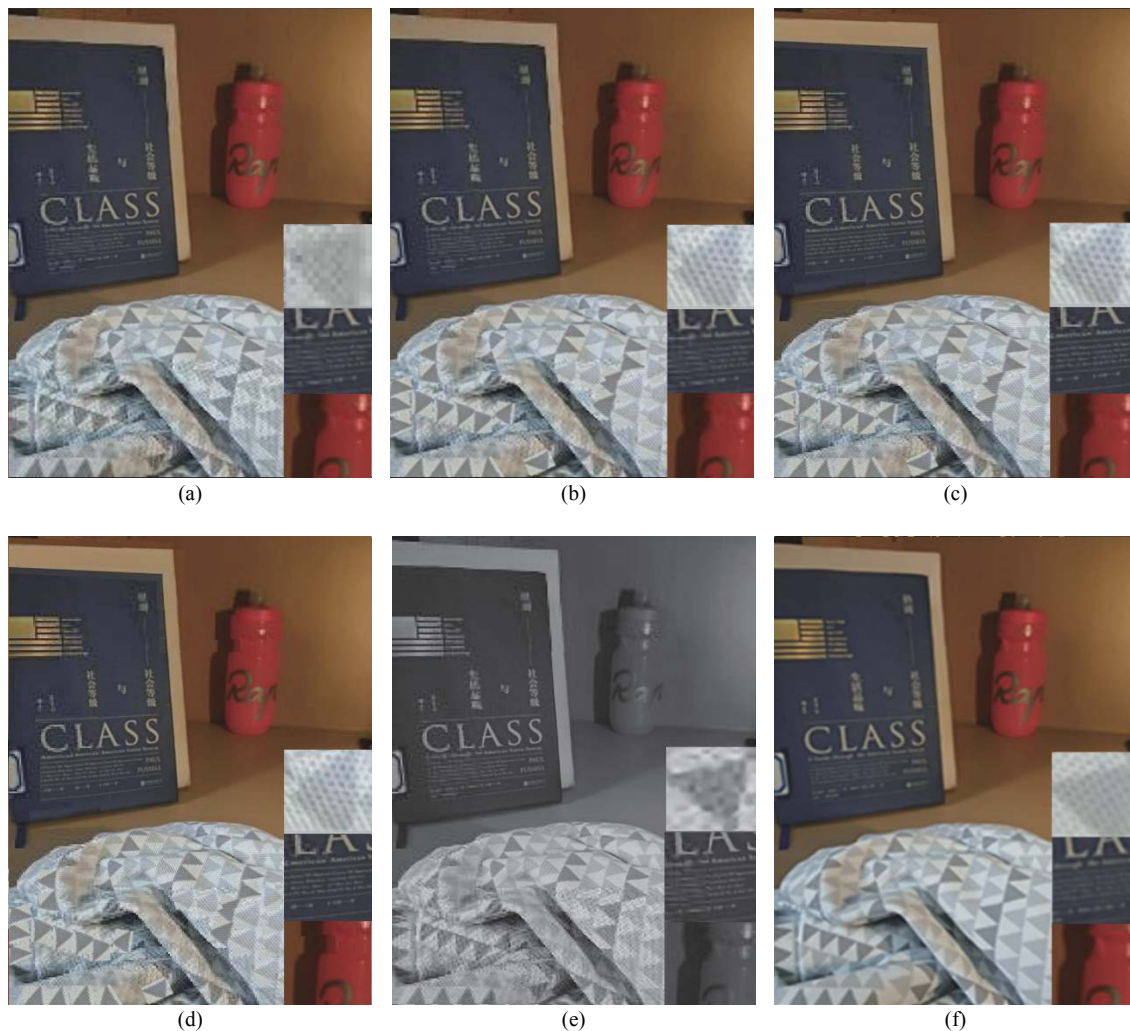


图 6 实拍图像基于对焦清晰度的大尺寸块匹配迁移结果

Fig.6 Real shot image focus based large size block matching migration results

表 2 不同算法的图像处理结果对比

Tab.2 Comparison of image processing results of different algorithms

Image evaluation function	Image processing algorithms 500×500					
	Label(f)	Paste(a)	SRCNN(e)	In the paper(b)	In the paper(c)	In the paper(d)
PSNR		25.85	27.40	24.47	24.21	24.19
SSIM		0.8665	0.736 4	0.817 3	0.812 6	0.811 0
MSE		169.08	118.28	232.31	246.41	247.65
Laplace	379.16	263.50	289.49	278.16	278.44	278.12

2.4 基于对焦清晰度的双分辨算法计算开销

空间相机等成像设备成像速度是远远快于图像处理的处理速度的。略微加长相机对焦时间和成像系统的存储空间负载,可以大大降低图像融合的计算负载,从而达到实时的图像处理。由表 3 可知,文中的算法是在相对计算时间和峰值内存开销上都远远

胜过单纯的 SRCNN,相对计算时间下降了约 60%,峰值内存开销约 35%。使用相对计算时间是多次程序运行时间的平均取整后得到。峰值内存开销以关闭所有前台运行程序后 2.7 GB/16 GB 作为计算基准,找到运行过程当中最大的内存开销然后减去基准值获得。可见,文中算法对于其他更加先进的深度学习

RefSR 改进到双分辨变焦算法当中还有巨大的提升作用。

表 3 不同算法的图像计算时间与内存开销

Tab.3 Image calculation time and memory overhead of different algorithms

Image	Relative calculation time				Peak memory overhead/GB			
	BI	SRCNN	ours	ours+	BI	SRNN	ours	ours+
4 000×3 000×1	2	33	14		1.0	13.2	8.3	
500×500×3	1	16	6	11	1.0	11.7	7.9	7.9

Note: “ours+” is the algorithm of this paper using multiple depth matrix, such as the paper(c) and paper(d) in Tab.2。

3 结 论

文中提出了双分辨相机中对焦清晰度的深度信息提取方法并讨论硬件和软件的可行性, 提出利用双分辨率多组对焦图像并提取深度信息的方案。文中结合深度信息基于小块匹配选择多种精度和匹配标准来修复双分辨率图像, 并将结果和 SRCNN 做了比较。论证了将深度信息匹配到深度学习的有参考图像深度学习网络中可行性, 大大提高了算法的效果和速度。

文中提出了结合深度信息来选择双分辨变焦算法中的参数的新方法。在不影响正常成像的条件下, 大大降低了开销, 提高了信息可靠度, 提高了信息量。

在此次的试验中发现, 深度图像精度对于文中算法在其他图像处理算法应用效果起到了决定性的作用。文中下一步的工作目标是使用双分辨相机读取相机马达的对焦参数, 并在相机对焦过程当中底层进行画面裁切配准和深度信息抓取, 进行工程上的测试, 从而提高算法的的准确性和稳定性。文中仅仅讨论了该方法应用在传统匹配方法中的效果, 相对于普通深度学习效果不错。希望可以将该方法用于最新的 RefSR 深度学习算法中并移植到双分辨成像当中来提高效果。

参考文献:

[1] He Guiran, Li Qi, Feng Huajun, et al. Continuous digital zoom of bifocal camera based on CNN feature extraction [J]. *Journal of*

Zhejiang University (Engineering Edition), 2019, 53(6): 1182–1189. (in Chinese)

[2] Ma Yuyu, Xu Zhihai, Feng Huajun, et al. Digital zoom algorithm for bifocal images based on high-frequency information reconstruction[D]. Hangzhou: Zhejiang University, 2018. (in Chinese)

[3] Chao Dong, Chen Change Loy, Kaiming He, et al. Learning a deep convolutional network for image super-resolution[C]//European Conference on Computer Vision (ECCV), 2014: 184-199.

[4] Zheng Haitian, Ji Mengqi, Wang Haoqian, et al. CrossNet: An end-to-end reference-based super resolution network using cross-scale warping[C]//European Conference on Computer Vision (ECCV), 2018.

[5] Zhang Zhifei, Wang Zhaowen, Lin Zhe, et al. Image super-resolution by neural texture transfer[C]//IEEE Conference on Computer Vision and Pattern Recognition(CVPR), 2019.

[6] Lei Dengfeng, Zheng Qunhui. An accurate camera depth of field calculation method [J]. *Information Technology*, 2013(8): 133–135. (in Chinese)

[7] Xu Yan, Wang Penghui, Liu Wei, et al. Comparative study of SIFT and SURF algorithms based on Matlab in image registration of drones [J]. *Mine Survey*, 2017, 45(6): 36–39. (in Chinese)

[8] Yin Lihua, Li Fanming, Liu Shijian. Panoramic image stabilization algorithm based on region segmentation and fusion [J]. *Infrared and Laser Engineering*, 2018, 47(9): 0926004. (in Chinese)

[9] Li Qi, Feng Huajun, Xu Zhihai, et al. Research on digital image sharpness evaluation function [J]. *Journal of Photonics*, 2002, 31(6): 736–738. (in Chinese)