PRELIMINARY INVESTIGATION ON POSSIBILITY OF SUPER RESOLUTION OF UAV ORTHOIMAGES

R. Matsuoka [a, b, *], K. Fukue [b]

[a] Geospatial Technology Department, Kokusai Kogyo Co., Ltd., 2-24-1 Harumi-cho, Fuchu-shi, Tokyo 183-0057, Japan -
ryuji_matsuoka@kk-grp.jp
[b] Tokai University Research & Information Center, 2-3-23 Takanawa, Minato-ku, Tokyo 108-8619, Japan -
ryuji@yoyogi.ycc.u-tokai.ac.jp, fku@keyaki.cc.u-tokai.ac.jp

Commission I, ICWG I/II

KEY WORDS: Super resolution, UAV, Orthoimage, Image processing, Overlapping, Sub-pixel, Pixel shift, Registration

ABSTRACT:

Since inspection of infrastructures using UAV images seems to be efficient, many systems for infrastructure maintenance using UAV images have been developed recently. For the purpose of more efficient image acquisition, we started an investigation on the possibility of super resolution (SR) of UAV images to obtain high-resolution (HR) orthoimages suitable for infrastructure maintenance. This paper reports an preliminary investigation using existing UAV images acquired for 3D measurement that were not be intended to be utilized for SR. We produced HR orthoimages by three SR methods: image interpolation of a single low-resolution (LR) image by cubic convolution, SR by resampling, and SR based on observation equations. Both SR by resampling and SR based on observation equations utilize multiple overlapping LR images. Results of the investigation demonstrate that SR based on observation equations using multiple overlapping images would be able to provide higher resolution orthoimages than those produced by an ordinary method. The results show that an inaccurate DSM utilized in SR processing degrades the quality of SR results as well. Furthermore, the results illustrate that the quality of the result of SR processing depends rather upon the characteristic of a lens utilized in image acquisition. We think that further investigations on SR using UAV images would be necessary in order to put SR to practical use.

1. INTRODUCTION

1.1 Inspection of infrastructures using UAV images

Japan has been building many infrastructures since the end of World War II. In the 21st century the importance of infrastructure maintenance is rapidly growing, and becomes higher than that of infrastructure building now. Currently, efficient inspection of infrastructures is seriously required because of a shortage of workers, especially skilled workers. Therefore, inspection of infrastructures using digital camera images is currently required in order to cut down on human works.

Inspection of infrastructures using images acquired by a digital camera on board an unmanned aerial vehicle (UAV), which becomes more popular recently, seems to be a promising means for efficient inspection. Many systems for maintenance of a concrete structure by using UAV images have been developed. Many of them is aimed to detect only hair cracks of sub-millimetre wide on a surface of a concrete structure.

A system developed by Mizukami, *et al.* (2018) aimed to detect cracks of 0.2 mm wide on a surface automatically by image processing techniques. Moreover, their system aimed to detect exfoliations of 10 mm deep on a surface, openings of 10 mm wide in a joint, and level differences of 10 mm high in a joint by using 3D measurement results provided by a piece of structure from motion (SfM) software. The performance of their system satisfies the standards of maintenance of such concrete structures as dams, head works, and open channels in Japan.

Their system was designed to acquire high-resolution images from 0.2 mm to 0.5 mm GSD (ground sampling distance) in order to detect cracks of 0.2 mm wide on a surface. On the other hand,

their system was designed to acquire images of equal to or smaller than 5mm GSD in order to detect exfoliations of 10 mm deep on a surface, openings of 10 mm wide in a joint, and level differences of 10 mm high in a joint.

Since the difference of GSD between two requirements is very large, their system utilizes two different sets of a camera and a lens. They adopted a lens of 135mm focal length for detection of cracks of 0.2 mm wide on a surface, and a lens of 50mm focal length for 3D measurement.

1.2 Need of super resolution of UAV images

Maintenance of such concrete structures as dams, head works, and open channels requires detection of cracks of 0.2 mm wide on a surface, while maintenance of such concrete structures as coastal dikes requires detection of cracks of simply 1.0 mm wide on a surface.

We considered that some methods to make resolution of observed images higher might enable to detect cracks of 1.0 mm wide on a surface by using UAV images of 5mm GSD. If UAV images acquired for 3D measurement by using a lens of 50mm focal length can be utilized for detection of cracks of 1.0 mm wide, the cost of inspection will be reduced.

Accordingly, we decided to start an investigation on possibility of super resolution (SR) of UAV images. Most cases of UAV photogrammetry in Japan adopt high overlapping ratios in image acquisition planning. One of the typical image acquisition plans is that the overlapping ratio of successive images along a flight strip, which is called OL for short, is 90%, and that of adjacent flight strips, which is called SL for short, is 60%. The high

overlapping ratios make a region of interest (ROI) be photographed by many UAV images.

Figure 1 shows a ROI, which 27 UAV images photographed as Figure 2 shows. This suggests that SR using multiple overlapping images may be applicable to obtain HR images. However, it is expected that an unstable flight of a UAV might make SR rather difficult.
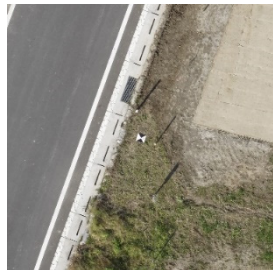

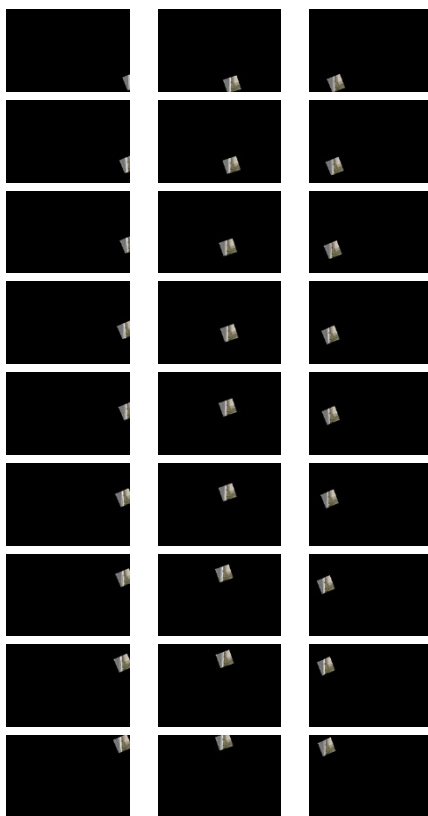Figure 1. Orthoimage of a ROI

Figure 2. UAV images photographing the ROI in Figure 1

Since infrastructure inspection using UAV images utilizes orthoimages frequently, our study focuses on the image resolution of UAV orthoimages. Considering that UAV images utilized for 3D measurement are highly overlapped, our study focuses on SR using multiple overlapping images as well. This paper reports our preliminary investigation by using existing UAV images acquired for 3D measurement which were not be intended to be utilized for SR.

2. SUPER RESOLUTION

2.1 Techniques of super resolution

Super resolution (SR) is to get resolution beyond the diffraction limit of an optical device. SR can be classified into two types:

one is a technique to obtain higher resolution during observation by using a dedicated observation system, and the other is a technique to obtain higher resolution after observation by using digital image processing. The former is called hardware-based SR, or device-dependent SR. The latter is called SR by data processing, digital SR, or mathematical SR. Since it is difficult to have a dedicated device for SR on board our UAV system, we decided to investigate methods of SR by data processing.

SR by data processing can be classified into two types: one is processed in the frequency domain, and the other is in the spatial domain (Yang, Huang, 2017). Since several SR methods in the spatial domain seems to be easily adaptable to UAV imaging system for inspection of facilities made of concrete, we decided to investigate these methods.

Furthermore, SR methods in the spatial domain can be classified into two types: one is methods to obtain a high-resolution (HR) image from a single observed low-resolution (LR) image, and the other is those to construct a HR image from several observed LR images. The former, which is denoted SR-S for short in the paper, is also called image interpolation (Meijering, 2002). We investigated two types of cubic convolution popular in image processing as SR-S.

2.2 Sub-pixel shift and overlapping image

Let $g(x, y)$ be an analogue image formed on the focal plane of an imaging system, $f(x, y)$ be a HR (fine) image, and $\{c_i(x, y)\}$ be observed LR (coarse) images. Here, n is the number of the observed LR images, and i $(i = 1, 2, \cdots, n)$ represents a sequence of multiple LR images. $f(x, y)$ and $\{c_i(x, y)\}$ are both sampled images.

$\{c_i(x, y)\}$ is obtained by shifting the sampling position by an amount Δ smaller than the pixel interval of the LR image when sampling $g(x, y)$. $\{c_i(x, y)\}$ is sometimes called a sub-pixel shift and overlapping image. The pixel shifts Δ may be unequally spaced when a camera on board a UAV acquires images. LR images utilized in SR are sub-pixel shift and overlapping images.
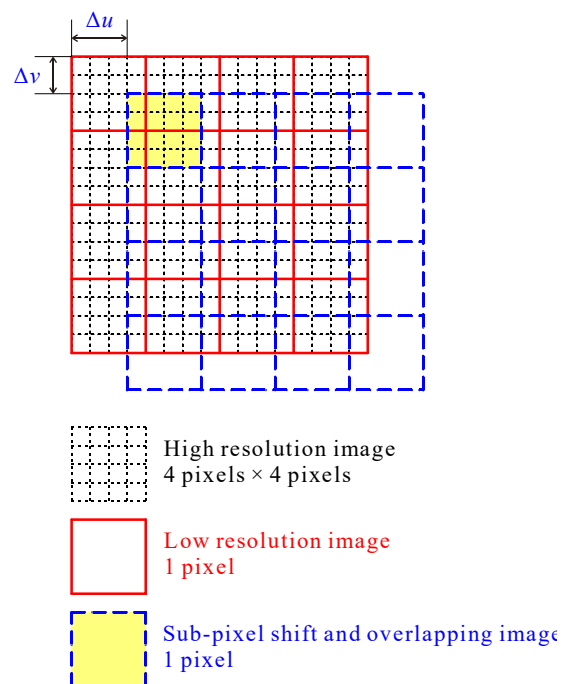

Figure 3. Sub-pixel shift and overlapping image

2.3 Super resolution by using multiple low resolution images

Popular methods of SR by using multiple LR images are (i) SR by resampling and (ii) SR based on observation equations.

2.3.1   SR by resampling: SR by resampling, which is denoted SR-R for short, is a method that constructs a HR image by averaging multiple resampled LR images. Each observed pixel of a LR image $\{c_i(x,y)\}$ is mapped to corresponding sampling position on a HR image $f(x,y)$ by a geometric transformation T_i. An average of the resampled LR images $\{c_i'(x,y)\}$ becomes a HR image $f(x,y)$.

$$f(x,y) = \frac{1}{n}\sum_{i=0}^{n}\{c_i'(x,y)\} \qquad (1)$$

Translation, Helmert transformation, affine transformation, or perspective transformation is often utilized as T_i depending on a problem. Nearest-neighbour interpolation or cubic convolution is popular as an interpolation method utilized in this method. This method is sometimes called SR by registration.

2.3.2   SR based on observation equations: A simple model that a LR image $\{c_i(x,y)\}$ is a sub-pixel shift and overlapping image is usually adopted in SR based on observation equations, which is denoted SR-O for short. We assume that each pixel of a HR image $f(x,y)$ and a LR image $\{c_i(x,y)\}$ is a square area, and is sampled at uniform intervals on an analogue image $g(x,y)$.

Let $h(u,v)$ be a point spread function (PSF), then a LR image $c_{Sx,Sy}(x,y)$ of a pixel shift (S_x, S_y) can be represented by the following equation:

$$c_{Sx,Sy}(x,y) = \int\int h(x+S_x-u, y+S_y-v)f(u,v)dudv \qquad (2)$$

If the grey level $g(x,y)$ of the analogue image takes a constant value $f(x,y)$ within the pixel of the corresponding HR image, that is to say the PSF is a square wave, $c_{Sx,Sy}(x,y)$ can be represented by the following equation:

$$c_{Sx,Sy}(x,y) = \sum_v\sum_u a(u,v)f(u,v) \qquad (3)$$

where $a(u,v)$ is the area of $f(x,y)$ included in the pixel (x,y) of $c_{Sx,Sy}(x,y)$ when the area of one pixel of $f(u,v)$ is 1.

Observation Equations (2) as to all pixels of all LR images $c_{Sx,Sy}(x,y)$ provide $f(x,y)$ if the number of observation equations is equal to or more than the number of pixels of the estimated HR image. Solving the observation equations is usually executed by iteration.

We adopt the iterative back projection (IBP) method proposed by Irani and Peleg (Irani, Peleg, 1993) in the investigation. Since the method is basic and simple but useful, it is one of the most popular method.

SR by the Irani and Peleg method as Figure 4 shows is performed by repeating the following processing.

Step 1:   Set an initial value $f^{(t=0)}(x,y)$ of the HR image. Any image can be utilized because it does not affect the convergence result. However, from the viewpoint of the convergence speed, the average value image of LR images acquired by using Equation (1) is often utilized.
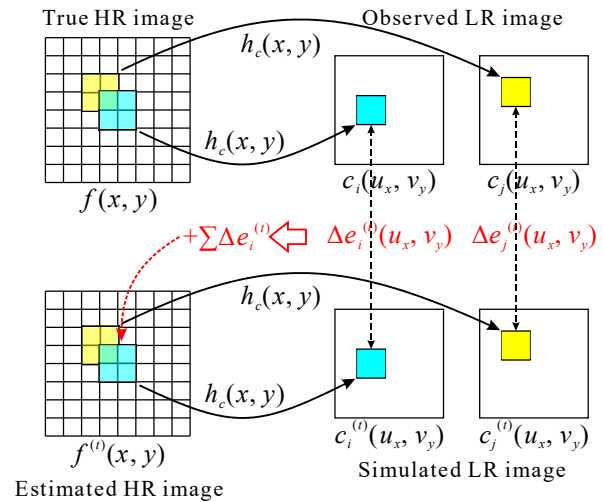

Figure 4. SR by the Irani and Peleg method

Step 2:   Perform the simulation of a LR image observation on $f^{(t)}(x,y)$. Let the obtained LR image be $\{c_i^{(t)}(u,v)\}$, then the observation simulation is expressed by the following equation.

$$c_i^{(t)}(u,v) = \sum_y\sum_x f^{(t)}(x,y)h_c(x_u-x, y_v-y) \quad (4)$$

where $h_c(x,y)$ is a PSF in the observation, and (x_u, y_v) is the coordinates on the (x,y) plane (HR image) corresponding to the centre of the pixel (u,v) on the LR image is located.

Step 3:   If $f^{(t)}(x,y)$ is a true HR image $f(x,y)$, $\{c_i^{(t)}(u,v)\}$ should be equal to the actual observed LR image $\{c_i(u,v)\}$. Let the observation simulation error $E^{(t)}$ be defined by the following equation:

$$E^{(t)} = \sqrt{\sum_i\sum_v\sum_u\left\{c_i^{(t)}(u,v) - c_i(u,v)\right\}^2} \qquad (5)$$

If $E^{(t)}$ is equal to or smaller than an appropriate designated value T_e, the processing is terminated.

Step 4:   Modify $f^{(t)}(x,y)$ so that $E^{(t)}$ becomes smaller. Then set the modified $f^{(t)}(x,y)$ as $f^{(t+1)}(x,y)$, and return to Step 2. Let (u_x, v_y) be the pixel of the LR image that observes the pixel (x,y) of the HR image, and $\Delta e^{(t)}(u_x, v_y)$ be the correction amount for a pixel (u_x, v_y). Moreover, let $w(u_x, v_y)$ be a weight as to $\Delta e^{(t)}(u_x, v_y)$, and its weighted average be a total correction amount $\Delta e^{(t)}(x,y)$ as to the pixel (x,y) of the HR image. $f^{(t+1)}(x,y)$ is represented by the following equation.

$$f^{(t+1)}(x,y) = f^{(t)}(x,y) + \Delta e^{(t)}(x,y) \qquad (6)$$

where,

$$\Delta e^{(t)}(x,y) = \frac{\sum_{u_x}\sum_{v_y}w(u_x,v_y)\Delta e^{(t)}(u_x,v_y)}{\sum_{u_x}\sum_{v_y}w(u_x,v_y)} \qquad (7)$$

$$w(u_x, v_y) = h_c(u_x, v_y) \qquad (8)$$

$$\Delta e^{(t)}(u_x, v_y) = \left\{c_i(u,v) - c_i^{(t)}(u,v)\right\}\frac{h_c(u_x,v_y)}{\alpha} \quad (9)$$

α is an appropriate scaling parameter.

Many studies on SR by using multiple LR images adopted the IBP method proposed by Irani and Peleg or its modified methods. Many of them utilized synthesize images and a few studies processed real images. Fukue *et al.* (2009) studied on the impact of radiometric density noise and geometric position noise. They utilized a digitized aerial photogram as a HR image, but LR images were synthesized by assumed observation equations. Matsumoto *et al.* (2009) reported an experiment on SR for detection of hair cracks of sub-millimetre wide on a concrete structure. They processed LR images acquired by using a real camera in not actual situation. There are few studies on SR by using multiple UAV images for inspection of infrastructure.

3. PROCESS OF SUPER RESOLUTION IN THE STUDY

We conducted SR processing in the study as follows:

[1] Estimate internal and external camera parameters of UAV images by using SfM software Pix4Dmapper Pro. This step has already been conducted by the dataset providers.

[2] Generate point clouds by using Pix4Dmapper. This step has already been conducted by the dataset providers as well.

[3] Create a digital surface model (DSM) from the generated point clouds.

[4] Create HR orthoimages by using the estimated internal and external camera parameters, and the created DSM. We adopt four interpolation methods: nearest neighbour interpolation (NN), bi-linear interpolation (BL), cubic convolution proposed by Riffman (1973) (CR), and cubic convolution proposed by Simon (1975) (CS). Results by CR and CS are utilized for evaluation of SR by using a single LR image (SR-S).

The interpolation functions *w(t)* of cubic convolution by Riffman and by Simon are represented by Equation (10) and Equation (11) respectively.

$$W(t) = \begin{cases} |t|^3 - 2|t|^2 + 1, & 0 \le |t| < 1 \\ -|t|^3 + 5|t|^2 - 8|t| + 4, & 1 \le |t| < 2 \\ 0, & 2 \le |t| \end{cases} \quad (10)$$

$$W(t) = \begin{cases} \frac{3}{2}|t|^3 - \frac{5}{2}|t|^2 + 1, & 0 \le |t| < 1 \\ -\frac{1}{2}|t|^3 + \frac{5}{2}|t|^2 - 4|t| + 2, & 1 \le |t| < 2 \\ 0, & 2 \le |t| \end{cases} \quad (11)$$

[5] Conduct SR by resampling (SR-R). Since the accuracy of camera parameters of each image and point clouds obtained by Pix4Dmapper are insufficient for SR by using multiple LR images, fine image registration is executed in the study. An image which photographed the ROI at the nearest to its image centre is selected as a reference orthoimage of the fine registration. Grey level of each orthoimage except the reference orthoimage are adjusted to that of the reference orthoimage as much as possible.

We adopt translation as geometric transformation, and we estimated shifts between orthoimages by least squares matching. Then each LR orthoimage for SR is created by using its estimated shift to the reference orthoimage. Results

by using four interpolation methods NN, BL, CR, and CS are utilized for evaluation of SR-R.

[6] Conduct SR based on observation equations (SR-O). We conduct fine image registration in the same way as the above-mentioned SR-R.

We adopt two types of a PSF $h_c(x, y)$: one is a square wave function, and the other is based on the Gaussian function. We assume that the latter has an effect on the 8-neighbourhood Ω of the target pixel (x_u, y_v) in the HR image, and is defined by the following equation.

$$h_c(u, v) = \begin{cases} \frac{1}{T} \exp\left(-\frac{u^2+v^2}{2\sigma^2}\right), & \text{inside } \Omega \\ 0, & \text{otherwise} \end{cases} \quad (12)$$

where,

$$T = \sum_{(u,v)\in\Omega} \exp\left(-\frac{u^2+v^2}{2\sigma^2}\right) \quad (13)$$

4. DATASETS UTILIZED IN THE STUDY

We utilized two existing datasets in the preliminary investigation: one is a dataset for a feasibility study of UAV photogrammetry for "i-Construction" (N. Takahashi, *et al.*, 2017), and the other is a dataset acquired in an evaluation experiment of the system developed by Mizukami, *et al.* The former is called Dataset T, while the latter is called Dataset M.

4.1 Dataset T

In the 2016 fiscal year the Ministry of Land, Infrastructure, Transport and Tourism of Japan started a program integrating construction and ICT in earthwork and concrete placing. The program was named "i-Construction". Dataset T was acquired in a field experiment to investigate whether the procedures of UAV photogrammetry following the standards for "i-Construction" are feasible or not.

The standard of UAV photogrammetry for "i-Construction" includes some requirements in image acquisition planning and measurement accuracy (Ministry of Land, Infrastructure, Transport and Tourism of Japan, 2016). In image acquisition planning, GSD should be equal to or smaller than 10mm. As for overlapping ratios, OL should be equal to or greater than 90%, and SL should be equal to or greater than 60%.

The field experiment by Takahashi, *et al.* was executed at a reinforcement site of the river embankment of the Edo River. Sony α6000 and ZEISS Loxia 2.8/21 were utilized in the experiment. Table 1 shows the specifications of α6000 and Loxia 2.8/21. A UAV photographed an embankment of about 80m by 160m immediately after earthwork was done on the embankment.

Camera	Sony α6000
Image sensor	23.5 × 15.6mm^2 CMOS
RAW recording pixels	6,000 × 4,000
Pixel unit	3.9μm square
Quantization bits	14-bit
Lens	ZEISS Loxia 2.8/21
Focal length (35mm film equivalent)	21mm (32mm)
Diagonal angle of view	91°

Table 1. Specifications of camera and lens of Dataset T

We utilized 891 UAV images that acquired by an image acquisition plan of 10mm GSD, 90% OL, and 60% SL. Figure 5 shows an orthoimage produced by Pix4Dmapper. Pix4Dmapper reported that the average GSD is 9.7mm and the mean reprojection error is 0.169 pixels.

Figure 5. Orthoimage of Dataset T produced by Pix4Dmapper

4.2 Dataset M

Dataset M was obtained in an evaluation experiment of the system developed by Mizukami, *et al.* 3D measurement of their system is conducted basically following the standard of UAV photogrammetry for "i-Construction". However considering the required accuracy of 3D measurement their system requests that GSD of images should be equal to or smaller than 5mm.

The experiment was executed at a coastal dike of the Nishi-Kunisaki Coast. Sony α9 and ZEISS Loxia 2/50 were utilized for 3D measurement in the experiment. Table 2 shows the specifications of α9 and Loxia 2/50. A UAV photographed a coastal dike of 100 m long.

Camera	Sony α9
Image sensor	$35.6 \times 23.8mm^2$ CMOS
RAW recording pixels	$6,000 \times 4,000$
Pixel unit	5.9μm square
Quantization bits	14-bit
Lens	ZEISS Loxia 2/50
Focal length	50mm
Diagonal angle of view	47°

Table 2. Specifications of camera and lens of Dataset M

We utilized 579 UAV images that acquired by an image plan of 3.6mm GSD, 93% OL, and 63% SL. Figure 6 shows an orthoimage by Pix4Dmapper. Pix4Dmapper reported that the average GSD is 3.8mm and the mean reprojection error is 0.114 pixels.

Figure 6. Orthoimage of Dataset M produced by Pix4Dmapper

5. RESULTS AND DISCUSSION

As for Dataset T, we show orthoimages of 4mm GSD created from UAV images of 9.7mm GSD here. Orthoimages around a square target of 30cm side for a control point for 3D measurement are mainly shown. On the other hand, as for Dataset M, we show orthoimages of 1mm GSD created from UAV images of 3.8mm GSD here. Orthoimages around a square chart of 28cm side are mainly shown. The chart shown in Figure 7 was placed in order to check the resolution of acquired images. Widths of black and white stripes are 1mm, 2mm, …, 8mm. Contrasts of all the following result images are enhanced.
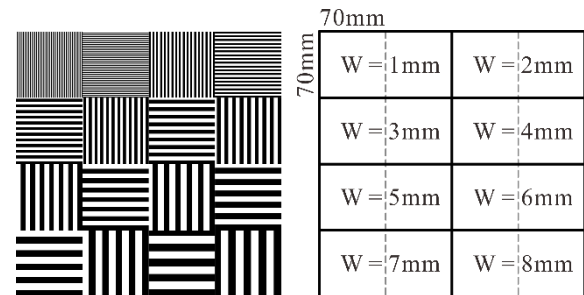

Figure 7. Chart for checking the resolution of acquired images

5.1 SR by using a single LR image (SR-S)

Figure 8 shows orthoimages created from the reference images which photographed the ROI at the nearest to its image centre by using four interpolation methods.

Jaggies can be found in the orthoimages created by using NN, while the orthoimages created by using BL are smoothed. Both

Interpolation	Dataset T	Dataset M
NN		
BL		
CR		
CS		

Figure 8. Results of SR-S

NN and BL cannot make HR images as expected. The results shown in Figure 8 indicate that CR and CS would not improve image resolution greatly. We cannot find the large difference between CR and CS as well.

5.2 SR by resampling (SR-R)

Figure 9 shows results of SR by resampling (SR-R) using overlapping LR images. The result orthoimages of Dataset T were created from 16 UAV images, and those of Dataset M were created from 40 UAV images.

Jaggies cannot be found in the orthoimages created by using NN, and the orthoimages created by using BL are not excessively smoothed. As for NN and BL, SR-R would be able to provide much higher resolution orthoimages than SR-S.

As for CR and CS, Figure 9 shows that SR-R would be able to provide better orthoimages than SR-S. However, we cannot find that CR and CS would improve the resolution of created orthoimages greatly. The small improvement is contrary to our expectation, and we suspect that its possible cause would be the degradation of resolution near to outer edges of an acquired image. We will discuss about it later.

As for interpolation method, CR and CS provide higher resolution orthoimages than NN and BL. However, we think that the differences of the quality of created orthoimages between four interpolation methods would not be so large. We cannot select a better interpolation method for SR-R between CR and CS the same as SR-S.

5.3 SR based on observation equations (SR-O)

Figure 10 and Figure 11 show results of SR based on observation equations (SR-O) using overlapping LR images. The results shown in Figure 10 were obtained when a square wave function was utilized as a PSF, while those shown in Figure 11 were obtained when CS was utilized as an interpolation method. The result orthoimages of Dataset T were created from 16 UAV images, and those of Dataset M were created from 40 UAV images the same as SR-R.

Figure 10 indicates that SR-O would be able to provide higher resolution orthoimages than SR-R as to all interpolation methods. However, the degree of the improvement of the resolution are small against our expectation. We suspect that the degradation of resolution near to outer edges of an acquired image would make the improvement of resolution smaller.

The differences of the quality of created orthoimages between four interpolation methods would be small. We recognize that CR and CS provide somewhat higher resolution orthoimages than NN and BL.

Figure 11 indicates that a square wave function or a Gaussian function of $\sigma = 3$ is desirable as a PSF. A Gaussian function of $\sigma = 1$ as a PSF would be unable to provide high quality
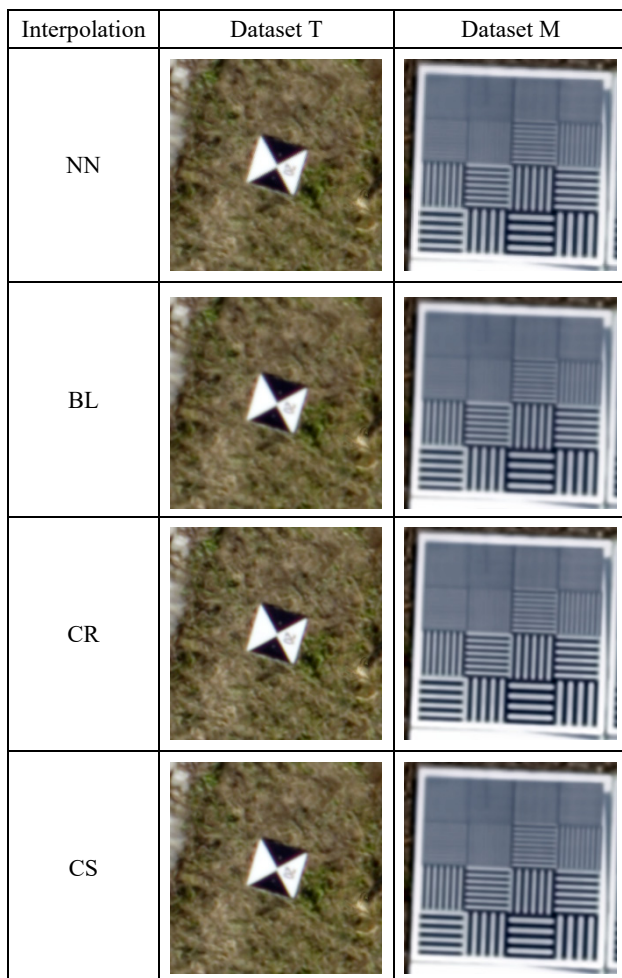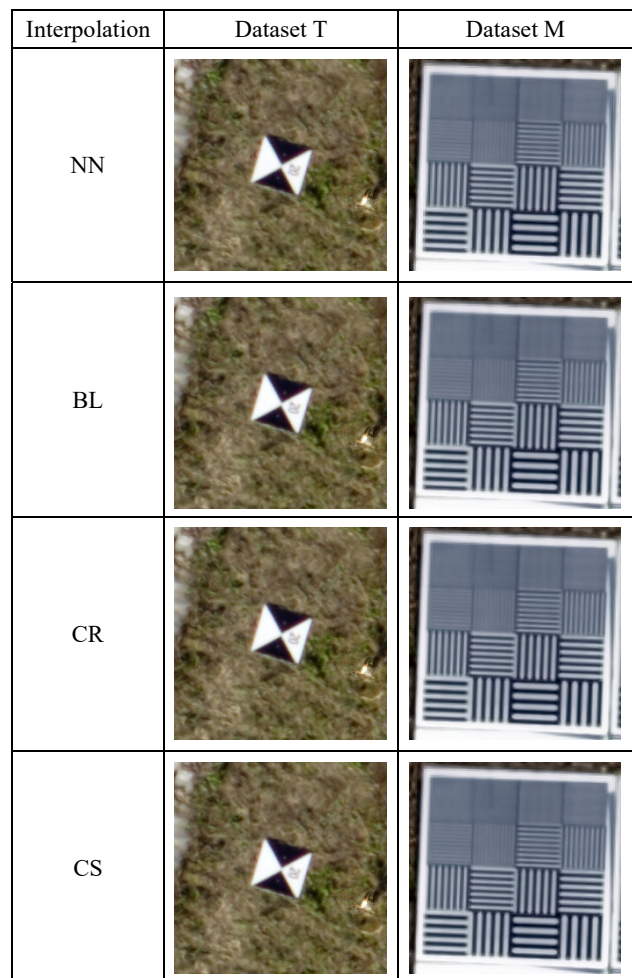

Figure 9. Results of SR-R

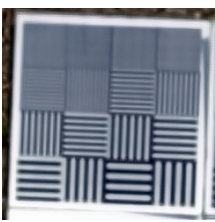Figure 10. Results of SR-O by using a square wave function

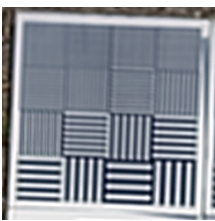| PSF	Dataset T	Dataset M
Square wave function		
Gaussian function $\sigma = 1$		
Gaussian function $\sigma = 2$		
Gaussian function $\sigma = 3$		

Figure 11. Results of SR-O by using CS

orthoimages. Since $\sigma = 3$ in the Gaussian function means that the degree of spreading (blurring) of a point object to adjacent pixels is extremely small, we think that the effect of a Gaussian function of $\sigma = 3$ as a PSF in SR-O is nearly equal to that of a square wave function as a PSF in SR-O.

Our current investigation results indicate that the best SR method of UAV orthoimages would be SR-O using a square wave function or a Gaussian function of $\sigma = 3$ as a PSF and CR or CS as an interpolation method. However, we recognize that the resolution of created orthoimages is not sufficient for our request.

5.4 Impact of an inaccurate DSM utilized in SR processing

Figure 12 shows two orthoimages of Dataset T. One is an orthoimage created from a reference image by using NN as an interpolation method, and the other is a result orthoimage of SR-O using a square wave function as a PSF and CS as an interpolation method.

The resolution of the orthoimage created by SR-O does not seem to provide higher resolution than that created from a single image at areas covered by grass. Moreover, a small vertical pole shown in the right centre part of the orthoimage created from a single image was disappeared in the orthoimage created by SR-O.

An inaccurate DSM utilized in SR processing caused these faults as Fukue *et al.* (2009) suggest. We should utilize an accurate

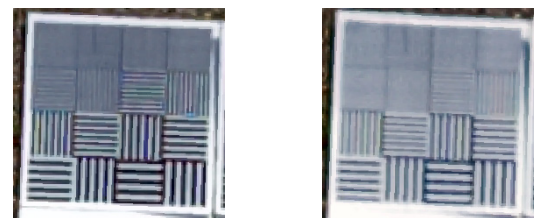DSM enough to provide HR orthoimages by SR-O. However, we consider that it would not be so easy.

(a) Orthoimage by using NN   (b) Result of SR-O

Figure 12. Orthoimage by using NN and result of SR-O

5.5 Impact of degradation of resolution near to outer edges

Figure 13 shows two orthoimages created from a single image of Dataset M by using NN. One is an orthoimage created from an image where a photographed ROI locates near to an image centre, and the other is one created from an image where a photographed ROI locates near to an outer edge of an image. Figure 13 indicates that the image resolution are not uniform all over an image.

(a) Near to an image centre   (b) Far from an image centre

Figure 13. Orthoimages created from a single image

Figure 14 shows a modulation transfer function (MTF) of ZEISS Loxia 2/50 (Carl Zeiss AG, 2017) that were utilized in image acquisition of Dataset M. Figure 14 indicate that the MTF of the lens depends on a radial distance on an image.
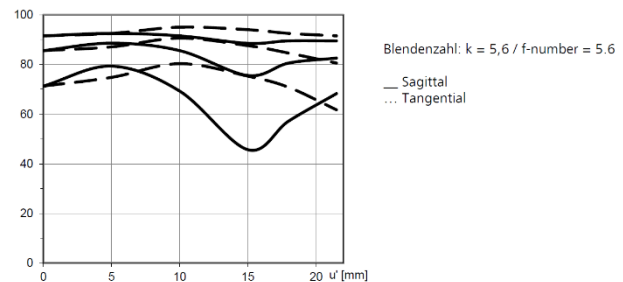
Figure 14. MTF of ZEISS Loxia 2/50

In order to investigate an impact of a degradation of resolution near to outer edges, we created two types of orthoimages by SR-O using a square wave function as a PSF and CS as an interpolation method. Let R be a radial distance from the image centre on a focal plane. One is orthoimages created from images where a photographed ROI locates inside the circle of $R = 18mm$, and the other is orthoimages created from images where a photographed ROI locates inside the circle of $R = 12mm$.

Figure 15 shows three orthoimages by SR-O: (b) using all 40 images, (c) 27 images where a photographed ROI locates inside the circle of $R = 18mm$, and (d) using 19 images where a photographed ROI locates inside the circle of $R = 12mm$.

The results indicates that exception of degraded images near to outer edges would be effective in creating HR orthoimages. The quality of orthoimages created from even half of acquired images would be superior to that of orthoimages created from all of acquired images.
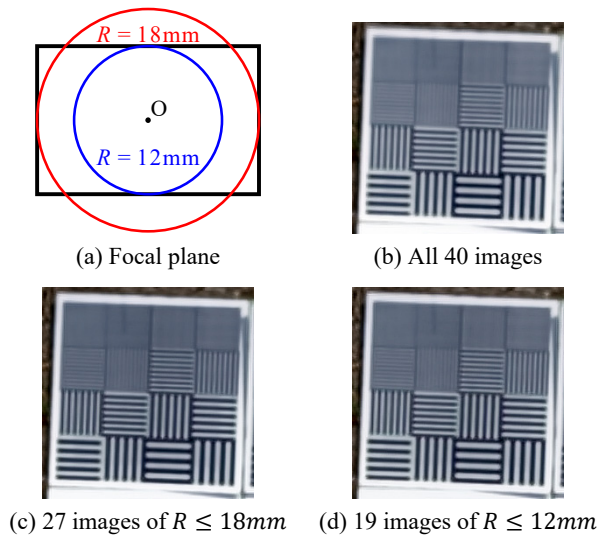

(a) Focal plane (b) All 40 images

(c) 27 images of $R \le 18mm$ (d) 19 images of $R \le 12mm$

Figure 15. Orthoimages created by SR-O

6. CONCLUSION

The results of our investigation demonstrate that SR-O would be able to provide higher resolution orthoimages than those produced by an ordinary method. On the other hand, the results shows that the improvement of resolution would not be large against the number of utilized overlapping LR.

The results show that an inaccurate DSM utilized in SR processing degrades the quality of SR results as well. However, we consider that it would not be easy to produce an accurate DSM corresponding to the required HR images. We have decided to develop a method to produce a more accurate DSM for SR processing.

Furthermore, the results illustrate that the quality of the results of SR processing depends rather upon the characteristic of a lens utilized in image acquisition. If a lens has a large degradation of resolution near to outer edges, one should not utilize acquired images that photographed a ROI near to outer edges.

We think that further investigations on SR using UAV images would be necessary in order to put SR to practical use. At first, we will conduct an experiment in order to design an image acquisition planning for not only 3D measurement but also SR processing.

REFERENCES

Carl Zeiss AG, 2017. ZEISS Loxia 2.0/50 12/16 Technical Specifications, https://www.zeiss.com/content/dam/consumer-products/downloads/photography/datasheets/en/loxia-lenses/datasheet-zeiss-loxia-250.pdf (17 May 2017)

Fukue, K., Sone, M., Sudo, N., Yokotsuka, H., Cho, K., Shimoda, H., 2009. Influence of Radiometric Density Noise and Geometric Position Noise in Reconstruction Based Superresolution, *Journal of the Japan Society of Photogrammetry and Remote Sensing*, vol. 48, no. 5, 289 - 298.

Irani, M., Peleg, S., 1991. Improving Resolution by Image Registration, *CVGIP: Graphical Models and Image Processing*, vol.53, no.3, 231 - 239.

Matsumoto, T., Sakagami, T., Sato, D., Kubo, S., 2009. Development of Crack Identification Technique by Super-resolution of Digital Images, *Proceedings of the Japan Concrete Institute*, vol. 31, no. 1, 2083 - 2088.

Meijering, E., 2002. A Chronology of Interpolation: From Ancient Astronomy to Modern Signal and Image Processing, *Proceedings of the IEEE*, vol. 90, no. 3, 319 - 342.

Ministry of Land, Infrastructure, Transport and Tourism of Japan, 2016. Draft of the Standard of Management System and Method of Finished Shape in Earthwork by UAV Photogrammetry.

Mizukami, K., Shiratani, E ., Kiri, H., Sekijima, K., Kaneko, T., Ooishi, T., Toyohuku, K., 2018. Automatic Deformation Extraction Method Using UAV for Efficient Sea Dike Inspection, *Journal of Japan Society of Civil Engineers*, vol. 74, no. 2, I-1435 - I-1440.

Rifman, S.S., 1973. Digital Rectification of ERTS Multispectral Imagery, *Proceedings of the Symposium on Significant Results Obtained from the ERTS-1 (NASA SP-327)*, vol. 1, sec. B, 1131 - 1142.

Simon, K.W., 1975. Digital Image Reconstruction and Resampling for Geometric Manipulation, *Proceedings of IEEE Symposium on Machine Processing of Remotely Sensed Data*, 3A-1 - 3A-11.

Takahashi, N., Wakutsu, R., Kato, T., Wakaizumi, T., Ooishi, T., Matsuoka, R., 2017. Experiment on UAV Photogrammetry and Terrestrial Laser Scanning for ICT-Integrated Construction. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLII-2-W6, 371 - 377.

Yang, J., Huang, T.S., 2017. Image Super-resolution: Historical Overview and Future Challenges, *Super-Resolution* edited by Peyman Milanfar. CRC Press, Boca Raton, 1 - 23.

ABBREVIATIONS

BL	bi-linear interpolation
CR	cubic convolution proposed by Riffman
CS	cubic convolution proposed by Simon
DSM	digital surface model
GSD	ground sampling distance
HR	high-resolution
IBP	iterative back projection
LR	low-resolution
MTF	modulation transfer function
NN	nearest neighbour interpolation
OL	overlapping ratio of successive images
PSF	point spread function
ROI	region of interest
SfM	structure from motion
SL	overlapping ratio of adjacent flight strips
SR	super resolution
SR-O	SR based on observation equations
SR-R	SR by resampling
SR-S	SR by using a single LR image
UAV	unmanned aerial vehicle