

# APPLICATION OF A HIGH-SPEED STEREOVISION SENSOR TO 3D SHAPE ACQUISITION OF FLYING BATS

Yijun Xiao, Robert B. Fisher

*School of Informatics, University of Edinburgh, 10 Crichton Street, Edinburgh, UK  
{yxiao,rbf}@inf.ed.ac.uk*

**Keywords:** High-speed video, 3D dynamic shape, stereo vision, motion, tracking.

**Abstract:** 3D shape acquisition of fast-moving objects is an emerging area with many potential applications. This paper presents a novel application of 3D acquisition for studying the dynamic external morphology of live bats in flight. The 3D acquisition technique is based on binocular stereovision. Two high-speed (500 fps) calibrated machine vision cameras are employed to capture intensity images from the bats simultaneously, and 3D shape information of the bats is derived from the stereo video recording. Since the high-speed stereovision system and the bat dynamic morphology study application are both novel, it was unknown to what extent the system could perform 3D acquisition of the bats' shapes. We carried out experiments to evaluate the performance of the system using artificial objects in various controlled conditions, and the knowledge gained helped us deploy the system in the on-site data acquisition. Our analysis of the real data demonstrates the feasibility of gathering 3D dynamic measurements on bats' bodies from a few selected feature points and the possibility of recovering dense 3D shapes of bat heads from the stereo video data acquired. Issues are revealed in the 3D shape recovery, most notably related to motion blur and occlusion.

## 1 INTRODUCTION

A relatively unexplored area of computer vision is the acquisition of 3D shapes from very fast-moving objects (Woodfill, 2004), despite a large amount of study in 3D shape recovery from various visual cues such as shading (Woodham, 1980), motion (Ullman, 1979), texture (Witkin, 1983), disparity (Trucco and Verri, 1998), active texture (Zhang and Huang, 2006), etc. This paper presents a novel application of high-speed 3D shape acquisition. The objective is to obtain 3D shape measurements of bats in flight when performing prey-hunting tasks, which are of biological and acoustic significance in answering questions related to bats' behaviour (Fenton, 2001).

A stereovision sensor has been employed for the shape acquisition. The reasons are twofold. First, high-speed video cameras are already available on the market with frame rates up to a few thousands fps (frames per second) at affordable cost (Pendley, 2003). Also the latest advances in stereovision sensing have been able to generate high quality range data (Siebert and Urquhart, 1994), even at video rate (Mark, 2006). However, 3D sensors based on other principles such as time-of-flight or active vision have not developed speeds that can

match video cameras. Therefore, we believe stereovision is a viable solution to high-speed 3D sensing in the near future. Second, stereo video recording provides valuable information to establish temporal correspondence on the 3D surface of the object, which may help study dynamic properties of the object (bat) shape.

This paper describes the progress in the bat observation application. The configuration of the stereovision sensor and the related performance evaluation are described in Section 2. Details of the sensor deployment in real data acquisition are given in Section 3, and Section 4 reports the analysis of the data we collected from live bats on-site.

## 2 SYSTEM EVALUATION

The stereovision system comprises two high-speed monochrome video cameras, two infrared (IR) light panels and two processing computers. The cameras have maximum resolution 1280x1024 and are running at 500 fps. Two sets of manual lenses with focal length 50mm and 75mm are employed, which

gives the stereovision sensor a working range [0.5m, 2m] in a window about 30cm(wide)x40cm(high).

We evaluated the sensor before the real data acquisition took place. First we examined the working range. We used a rigid planar surface with sharp texture as the ground truth object. The stereo images of the planar surface were acquired and then processed using DI3D™ stereo photogrammetry software [Khambay, 2008] to generate 3D images of the object. Each 3D image was fitted with a 3D plane, and the RMS (Root Mean Square) value of the fitting residuals was calculated to evaluate the accuracy of the 3D image against the ground truth.

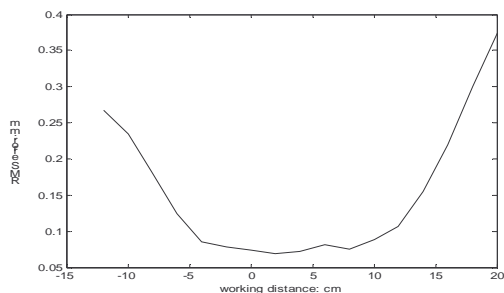


Figure 1: RMS errors of measuring a 3D plane around working distance 80cm using 50mm lenses.

We placed the planar surface at a number of positions and calculated corresponding RMS errors. Fig. 1 illustrates the RMS errors obtained from all test positions around working distance 80cm using 50mm lenses. It can be seen that the RMS errors exhibit clearly a basin shape with a flat bottom of about 10-15cm. The basin shape was observed in tests with other configurations of the stereovision system as well. We thought that the basin shape was caused by the depth of field of the lenses.

The planar object was also used to check spatial coherence of the sensor. A 3D image of the planar surface was partitioned into 30x30 patches. Each patch was fitted with a 3D plane and a RMS error was calculated. The RMS errors are shown in Fig.2, where pseudo colour represents the RMS errors. It appears the RMS errors are randomly distributed on the planar surface, which suggests no obvious bias to a specific spatial location in the 3D measurements.

We also conducted tests to reveal dynamic properties of the sensor. We used a well-textured ball as the ground-truth object. The object was swung over the sensor in three orthogonal directions (horizontal, vertical and depth). By fitting a 3D sphere to the 3D image of the ball, we can calculate the position of the centroid of the ball, which can be then used to estimate the ball speed. The fitting residuals can be used to calculate RMS errors. The

RMS errors related to motion in the depth direction are illustrated in Fig. 3. It can be seen the error curve at low speed (0.9m/s) exhibits a basin shape just like in the static test (Fig.1). At medium and high speed, the error curves are much more flattened and noisier, which reflect the influence of motion blur when speed gets higher (>3m/s). RMS errors with respect to horizontal and vertical motion in our tests are larger than those with respect to motion on the depth direction at the same magnitude. Especially the horizontal motion caused the highest level of RMS errors. The sensitivity of stereovision to motion parallel to epipolar line is hypothesized.

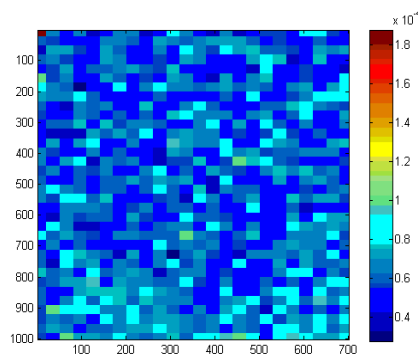


Figure 2: Spatial coherence: RMS errors of fitting planes to 30x30 blocks in a 3D image.

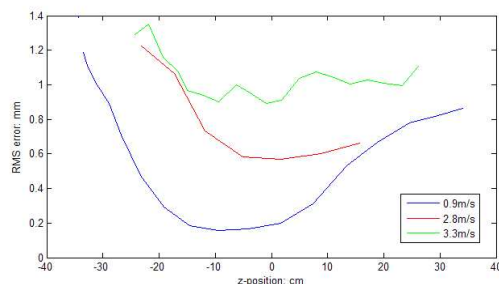


Figure 3: RMS errors of fitting a sphere to 3D images of a ball swung along the depth direction.

### 3 DATA ACQUISITION

The real data acquisition took place at two sites: Odense in Denmark and BCI (Barro Colorado Island) in Panama. Two groups (water-trawling and insect-gleaning) of bats were examined. For the water trawling bats, the stereovision system was placed in the water to capture the front view of bats flying towards the cameras (Fig. 4). The flying path

of bats was constrained in the depth direction of the sensor. Bait was put in the water at regular places. The bats usually flew through the narrow water path to take the bait, since they needed a distance to take off from the water surface. The bats were therefore trained to move in the depth direction of the stereovision sensor.

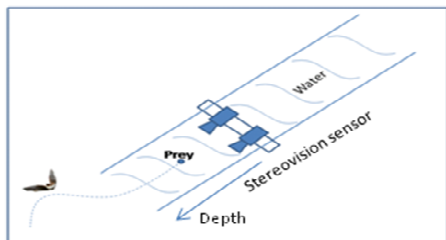


Figure 4: Sensor setup for water trawling bats.

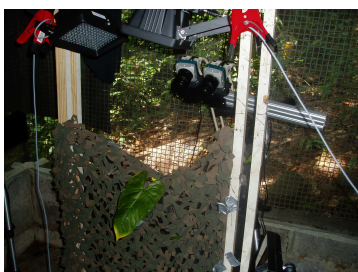


Figure 5: Sensor setup and for insect gleaning bats.

For the insect gleaning bats, object speed is not crucial in 3D acquisition, since the bats usually reduce their speeds (in our measurements the speed was below 1m/s) to hover around the prey in search of a good hunting position. To obtain the best image quality, we placed the sensor focussed at the closest point to the prey (Fig.5). The prey (insect) was placed on a leaf among a camouflage bush. The position of the prey on the leaf was adjustable to allow study of the bat's searching behaviour before hunting the prey. The focusing point of the stereovision sensor was on the centre of the leaf. The orientation of the leaf was adjusted almost perpendicular to the depth direction of the sensor to avoid observing occlusion by the leaf. With this setup, the sensor was able to capture larger foreground objects.

## 4 DATA ANALYSIS

The first analysis of the stereo data recording is to make some 3D measurements on the bats' bodies. A

few feature points (landmarks) were manually placed in one (reference) frame of the left and right image sequences. Once the landmarks were selected in the reference frame, they were tracked across the remaining frames of the image sequences using an algorithm based on optical flow (Ogale and Aloimonos, 2007). The tracking is currently running in a semi-automatic manner. In some frames where the scene has drastic changes, location prediction of the landmark points could be severely inaccurate, and then manual marking of the locations was involved.

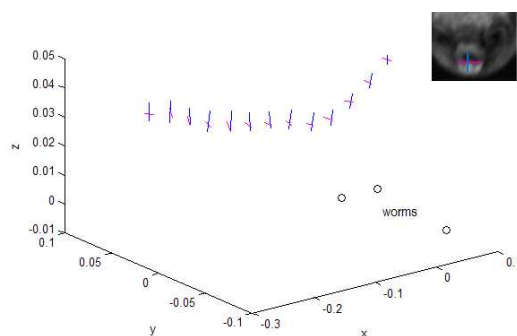


Figure 6: Mouth measurements of a bat in flight hunting prey.

Once the landmark points were extracted from the left and right image sequences, stereo triangulation was performed to calculate the corresponding 3D points, which can then be used to derive some measurements such as speeds, sizes, mouth openings of the bats, etc. Fig. 6 illustrates an example of such 3D measurements. The four corners of the mouth of the bat were measured in 64 consecutive frames. The recovered 3D positions of the mouth corners are depicted in Fig. 6, where the red lines represent the widths of the mouth and the blues represent the heights. The positions of the worms in Fig.6 were also recovered, from which the water surface was calculated. The 3D positions of the bat mouth corners and the worms in Fig. 6 are displayed in a coordinate system in which the water surface is aligned with the x-y plane. Such arrangement of coordinates renders the altitude of the bat in flight above the water level (z-coordinate) explicit. It is clearly seen that the bat approached the prey in a low and flattened path and then lifted above the water surface immediately after grabbing the prey.

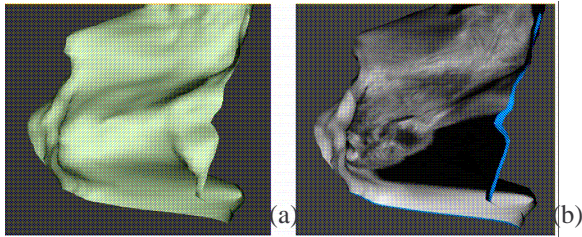


Figure 7: 3D shape of *M. daubentonii* recovered stereo images.

3D shape recovery from the stereo data was also investigated using the DI3D™ software. We considered three main issues which may affect the stereo software to obtain 3D shapes of bats such as: 1) lack of texture on bats' bodies; 2) occlusion of body parts; 3) motion blur. Our experiments so far have shed some light on these concerns. Firstly, it was found that the texture of bats could be revealed under proper illumination. Fig.7(b) shows a bat head's fine texture when the bat flew through the working range of the sensor on which the IR light was concentrated. With the texture visible to the cameras, the stereo software was able to recover part of the 3D geometry of the bat head as shown in Fig.7(a). Secondly, motion blur did take its toll on 3D shape recovery. For instance, the bat's head in Fig.7(a) was smoothed with some fine shape details missing. Thirdly, the effect of occlusion was evident. For instance, the bat's ears were squeezed onto the head in Fig.7. Despite these defects, the results show that it is possible to recover 3D shapes of bats using stereovision methods.

## 5 CONCLUSIONS

This paper reports an application of using a 500 fps stereovision sensor to capture 3D external morphology of bats in flight. We discussed the data acquisition scenarios and evaluated the performance of the sensor accordingly. Stereo data were acquired from four species of live bats, and the preliminary analysis of the data confirmed that it is feasible to obtain 3D shape information of bats in flight for the chosen species using the stereovision method at 500fps. A number of issues were revealed in 3D shape recovery related to motion blur and occlusion, which helps identify the problems we will be working on and revise the expectation of quality of 3D measurements we can draw from the stereo data.

## REFERENCES

- Fenton, M.B., 2001. *Bats*. Checkmark Books.
- Khambay, B., Narin, N., et al, 2008. Validation of reproducibility of a high-resolution three-dimensional facial imaging system, *British Journal of Oral and Maxillofacial Surgery*, 46(1):27-32.
- Ogale, A.S., Aloimonos, Y., 2007. A roadmap to the integration of early visual modules, *International Journal of Computer Vision*, 72(1):9-25.
- Pendley, G.J., 2003. High-speed imaging technology; yesterday, today & tomorrow, *Proceedings of SPIE; 25th International Congress on High-Speed Photography and Photonics*.
- Scharstein, D., Szeliski, R., 2002. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms, *International Journal of Computer Vision*, 47:7-42.
- Siebert, J.P., Urquhart, C.W., 1994. C3D: a novel vision-based 3-D data acquisition system, *Proceedings of the European workshop on combined real and synthetic image processing for broadcast video production*, Hamburg, Germany.
- Trucco, E., Verri, A., 1998. *Introductory techniques for 3-D computer vision*, Prentice Hall.
- Ullman, S., 1979. The interpretation of structure from motion, *Proceedings of the Royal Society of London B.*, 203, 405-426.
- Mark, W. van der, Gavrilu, D.M., 2006. Real-time dense stereo for intelligent vehicles, *IEEE Transactions on Intelligent Transport System*, 7(1): 38-50.
- Winderu, R.J., Darvannb, T.A., et al, 2008. Technical validation of the Di3D stereophotogrammetry surface imaging system, *British Journal of Oral and Maxillofacial Surgery*, 46(1):33-37.
- Witkin, A.P., 1981, Recovering surface shape and orientation from texture, *Artificial Intelligence*, 17(1): 17-45.
- Woodham, R.J., 1980. Photometric method for determining surface orientation from multiple images. *Optical Engineering*, 19(1): 139-144.
- Woodfill, J.I., Gordon, G., Buck, R., 2004. Tyzx DeepSea high speed stereo vision system, *Proceedings of Computer Vision and Pattern Recognition Workshop, CVPRW '04*. Washington, D.C., USA.
- Zhang, L., Curless, B., Seitz, S. M., 2003. Spacetime Stereo: Shape Recovery for Dynamic Scenes. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Madison, USA.
- Zhang, S., Huang. P. , 2006. High-resolution, real-time 3-D shape measurement, *Optical Engineering*, 45(12).