

# BeNoGo Image Volume Acquisition

Hynek Bakstein      Tomáš Pajdla      Daniel Večerka

## Abstract

This document deals with issues arising during acquisition of images for IBR used in the BeNoGo project. We describe the acquisition setup at our disposal and we focus on specific issues for each device as well as general issues of the acquisition process. We describe the main difficulties and we try to characterize an “ideal” scenario which can be acquired with our technology.

## 1 Introduction

We first review our acquisition setups and focus on the resolution they provide, the number of images required to capture a scene, and the time necessary for acquisition, based on the frame rate of the device. Then, we summarize characteristics of the visualization devices and compare them to the acquisition setups. Based on the characteristic of the acquisition devices with sufficient resolution, we can deduce some technical limitations of the capturing process, mainly due to a long acquisition time. Then, we focus on issues specific to the location, where the image are to be captured, itself. Finally, we propose three possible locations minimizing problems of the acquisition process. It should be noted that we focus on the technical issues only, not on other aspects of the location, such as attractiveness, natural scenario, etc. It would be ideal to find such a place, which minimizes the technical issues while maximizing the requirements of the measurement partners. We think that a location like the Technical museum is a good compromise.

## 2 Acquisition setups

Our acquisition setup has developed and improved over time, see Figure 1 for overview. First, we were using a digital video camera with limited field

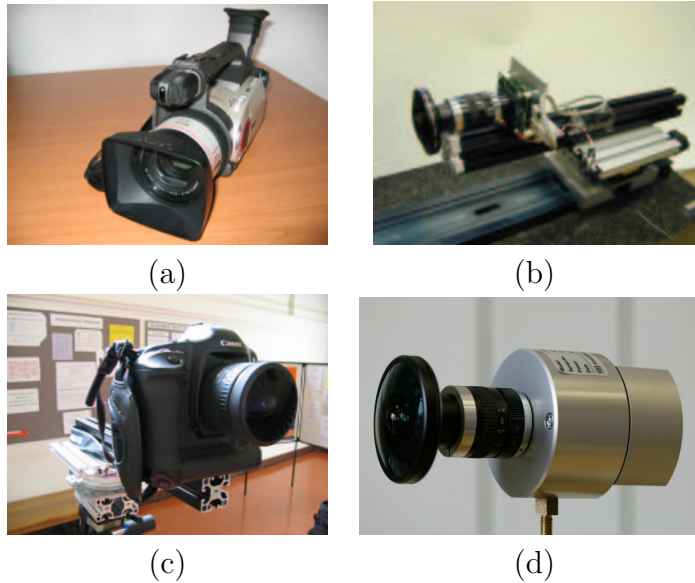


Figure 1: Acquisition setups. (a) Canon DV camcorder, (b) Pixelink camera with FC-E8 fish eye converter, (c) Canon 1Ds with Sigma 8mm lens, and (d) Vosskühler CCD-4000C with FC-E8.

of view (a). It was used just to get some initial data before better cameras arrived. On the other hand, it allowed for fast data acquisition at 25 frames per second (fps). Also, small FOV provided high number of pixel per one degree, 14.4.

Next generation of setups is the first omnidirectional setup consisting of a Pixelink Firewire camera with Nikon FC-E8 fish eye adapter (b). This setup provides 1Mpix images ( $1200 \times 1024$  pixels) at 9 fps. Such a resolution showed to be not sufficient for high quality display system, when pixelization was visible. Also the colour fidelity of images is poor, there is a lot of noise. On the other hand, decent frame rate allows fast data acquisition.

Presently, we use setup composed from a Canon 1Ds digital camera and Sigma 8mm fish eye lens (c). Canon 1Ds is a professional digital camera with 11Mpix resolution and high quality images. This resolution is sufficient for our imaging system, however, the amount of data is huge. The camera also needs 3 seconds to store the image either to a memory card or to a computer. This results in long acquisition times. This setup offers image quality superior to any other devices at our disposal in both image resolution and color fidelity.

Camera	Img. Resolution	FOV	Fps	Acq. time	# images
CanonDV	$720 \times 576$	$\approx 60$	25	3m	$\approx 5000$
Pixelink	$1200 \times 1024$	180	9	6m	$\approx 3000$
Canon 1Ds	$2560 \times 2560$	180	0.3	40m (5h)	$\approx 700(5000)$
VDS CCD-4000C	$2048 \times 2048$	180	5	2,5m (13m)	$\approx 700(4000)$

Table 1: Characteristics of our acquisition devices.

The newest setup at our disposal is a Vosskühler CCD-4000C camera equipped with a Nikon FC-E8 fish eye lens (d). This camera is capable of capturing 5 frames per second at 4Mpix resolution ( $2048 \times 2048$ ). This resolution is sufficient for many purposes. This is a very recent piece of equipment and we are now in a testing phase when we evaluate the noise level and color fidelity of the images. However, this is the most promising camera because of a combination of high frame rate and high resolution. For example, we should be able to capture several discs at one location during a single session. On the other hand, we have to deal with issues such as a manual white balance, huge data flow (20 MB/s), and portable data storage with sufficient capacity (about 180GB).

Characteristics of all three setups are listed in Table 1. We should note here that during IBR, omnidirectional mosaic images are created. These images cover  $180^\circ$  vertically and  $360^\circ$  horizontally. Each mosaic image is composed from input images acquired by the acquisition setup, processed into so called volume image [1, 2]. Here we just mention that the volume images have the same height in pixels as the input images. Thus, the vertical dimension of the mosaic image is given by this height. The horizontal dimension should be twice as big to maintain the same pixel per degree resolution in both horizontal and vertical direction. As a result, the number of images should be the same as the horizontal resolution in pixels. It is very difficult to handle so many number of images with the Canon 1Ds, where we would have to acquire 5000 images which would take about 5 hours.

Another problem with acquisition devices is that they do not allow to capture dynamic scenes. Even the fastest setups need some minutes to get the images. Moving people or other moving objects then create artifacts in images rendered by IBR, see Figure 2(a). Sun moving in the sky results in moving shadows, clouds result in different illumination of the scene in different images, depicted in Figure 2(b). These effects create stripes with

varying brightness in images from IBR. All these issues have to be solved. It is clear that faster acquisition time minimizes these effects.

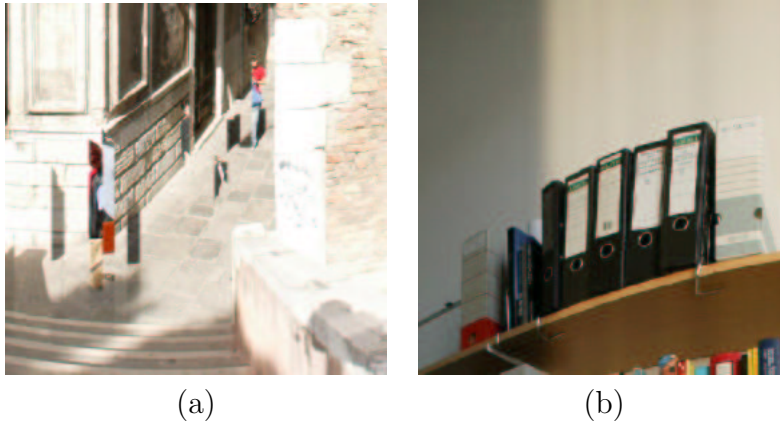


Figure 2: Issues arising from long acquisition times. (a) Artifacts from moving objects. (b) Different illumination caused by clouds.

### 3 Visualization devices

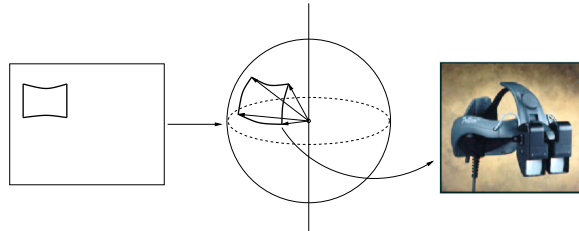


Figure 3: Viewing window simulates the limited field of view of the eyes of the viewer. Image of the HMD is from Kaiser Electro-Optics, Inc.

The viewer does not look at the whole omnidirectional images created by IBR, he is presented with some small part which we call *a viewing window*, see Figure 3. This smaller part correspond to limited FOV of the display device. Based on vertical resp. horizontal resolution in pixels  $r_v$  resp.  $r_h$  and FOV  $\varphi_v$  resp.  $\varphi_h$  of this device, we can compute the required resolution of

the mosaic images  $mr_v$  and  $mr_h$  resp. and hence the required resolution of the input images

$$mr_v = \frac{180^\circ}{\varphi_v} r_v, mr_h = \frac{360^\circ}{\varphi_h} r_h . \quad (1)$$

In Table 2, we summarize vertical and horizontal resolutions per degree of our acquisition setups and we presents also these resolution for two head mounted displays (HMD) used in our project. Note that Visette Pro has resolution of  $640 \times 480$  pixels but for all colors together. Thus, the effective resolution is smaller. We can conclude that Canon 1Ds and Vosskühler CCD-4000C provide images with resolution sufficient for displaying in both HMDs used in our project. Devices like Panorama or Cave might require even higher resolution. Images captured by Pixelink camera have resolution sufficient for display in Visette Pro.

Device	Vert. res.	Hor. res.	Vert. FOV	Pix. per deg.
CanonDV	576	768	40	14.4
Pixelink	1024	2048	180	5.7
Canon 1Ds	2560	5120	180	14.2
VDS CCD-4000C	2048	4096	180	11.4
Visette pro	160	213	45	3.5
V8	480	640	45	10.6

Table 2: Number of pixels per degree for our acquisition setups and some display devices.

We would like to conclude this section with a note that the number of images and thus the acquisition time can be reduced using some knowledge of the scene geometry. The better knowledge, the lower is the decrease of quality of the IBR images. More details can be found in [3]. However, the image number is still quite high and the acquisition time will be at least several minutes.

## 4 Scenario related issues of acquisition

Apart from the technical issues overviewed above, which are mainly results of the long acquisition time, there exist other limiting factors which play

important role for a scenario selection. These factors are related directly to the location of the scenario and include accessibility and repeatability of measurement and acquisition, control over the environment, and many others. We will summarize these factors in this section.

- **Accessibility** is a basic requirement for a scene. We have to be able to access the location. We should also be able to repeatedly reenter the location and it would be best to be alone in the location, without any visitors. Therefore, the best option is a space which can be closed for visitors or is not visited so frequently at least during one day in a week. We need some visitors, however, for benchmarking. This condition can be fulfilled by both indoor and outdoor locations.
- **Control** over the environment means the level of control over the lighting, placement of objects in the scene and so on. This also applies to long term control over the place, so that object and lights will be at the same place when we reenter the location after some (possibly long) period of time. This may be a problem with outdoor scenes, where season changes affect the environment considerably, for example, snow in winter, trees blossoming in spring, etc. We have not taken this into account in the viewpoint scenario.
- **Stationariness** of the location. No moving object, this includes also clouds, trees moving in the wind, etc. This may be again a problem with outdoor scenarios since there can be clouds, wind moving the trees, Sun moving in the sky, etc.

We conclude this session with a note about the REX limitations. Right now, we have experience with capturing images for a point and disc REX. The disc can have a diameter up to 60cm. The location should therefore take this into account. In the near future, we plan to test a line/plane REX.

## 5 Conclusion - proposal of an “ideal” environment

The above summarized issues point us to an indoor scenario in a place where we have repeated access and where the objects in the scene are not moved so

frequently. We propose three different indoor scenarios, two sharing similar characteristics, one outstanding.

Scenario 1 — **The technical museum**. This location has been captured several times and we have a possibility to enter the place when it is closed for public. Ideal lighting source is the glass ceiling, resulting in almost constant lighting condition during clear sky days. Moreover, the glass is matted, therefore the Sun is not visible in the images. Repeated benchmarking is possible, since the place is frequently visited.

Scenario 2 — **The botanical garden**. Even though we do not have access to the place while it's closed for public, it is not so frequently visited during workdays, and therefore we are able to acquire images without moving people. Direct sunlight should be avoided, thus the acquisition should be done during overcast weather. Other aspects of the environment are controlled and the place is visited frequently during weekends, which allows benchmarking. The management of the garden also cooperates with us, we have already acquired our initial scenario there.

Scenario 3 — **The office**. This scenario is not a public place, unlike the previous two. Thus, we have a total control over the place. It is ideal for the technical, B & C type, tests. We have already demonstrated capability of capturing more than 5000 images in this location. It should be noted, that we may move to another building in September, therefore we would have to use another office as a location. Anyway, this particular office was captured both during the night (full dataset, more than 5000 images) and during the day (sparse dataset).

## 6 Conclusion

We have summarized our acquisition setups and issues arising during acquisition procedure. We have analyzed both the technical aspect related to different setups as well as general factors depending on the scene. Finally, we have proposed three scenarios minimizing the difficulties of the acquisition procedure. All partners are encouraged to send us their comments, this document is open and will develop in time.

## References

- [1] Hynek Bakstein and Tomáš Pajdla. Ray space volume of omnidirectional 180x360 deg. images. In Ondřej Drbohlav, editor, *Computer Vision — CVWW'03 : Proceedings of the 8th Computer Vision Winter Workshop*, pages 39–44, Prague, Czech Republic, February 2003. Czech Pattern Recognition Society.
- [2] Hynek Bakstein and Tomáš Pajdla. Rendering novel views from a set of omnidirectional mosaic images. In *Proceedings of Omnivis 2003: Workshop on Omnidirectional Vision and Camera Networks*, page cdrom only, Los Alamitos, CA, June 2003. IEEE Press.
- [3] Hynek Bakstein and Tomáš Pajdla. Visual fidelity of image based rendering. In Danijel Skočaj, editor, *Proceedings of the Computer Vision Winter Workshop 2004 (CVWW'04)*, pages 139–148, Ljubljana, Slovenia, February 2004. Slovenian Pattern Recognition Society.