IMAGE PROCESSING CONSIDERATIONS FOR SIMPLE REAL-TIME RESTRICTOMETRIC FLUID-BASED USER INTERFACES

Steve Mann, http://eyetap.org

University of Toronto, Department of Electrical and Computer Engineering

ABSTRACT

The FUNtain (TM) is a new input device in which a user inputs data by direct interaction with fluid, such as with one or more air or water jets. There is usually a (re)strictometer for each jet that measures the degree of restriction of fluid flow for that jet, such that the time-varying changes in restrictometric quantities provide an input to another process. For example, 61 restrictometers are used to provide a direct-acting waterpipe organ as a musical instrument in which notes are played by blocking fluid (air or water) emerging from each of the 61 holes spaced along a copper pipe that forms the housing for the musical instrument. When used with liquids, such a fluid-user interface based on the image processing of water sprays is proposed to create a new form of input device. In particular, the methodology exploits the unique optical properties of fluids that become manifest when a fountain or spray is illuminated as it would typically be for an exhibit, display, or performance. Most notably, the form of illumination that is most desirable for aesthetic reasons turns out to assist simple image processing of the water flow. In this way, an interactive performance environment is created that is ideally suited to real-time image processing as a means of data input. The resulting methodology gives rise to a liquid-user-interface, played by one or more people. interacting with water jets, such as by blocking, partially blocking, diverting, or otherwise engaging with the spray.

1. INTRODUCTION

Image processing in combination with water is used in the Poseidon system [1], the Drowning Early Warning System (DEWS)[2], and Swimguard: three commercial efforts at providing computer-based drowning detection. Other work, such as particle image velocimetry, exists to, for example, monitor fluid flow in pipes using laser interferometry. Image processing has also been used to initiate fluid flow (e.g. to automate the process of flushing toilets, turning on faucets, etc.) [3]. However, no previous work has been done on using image processing for the creation of a user interface or input device that is based on tracking fluid flow.

Recent developments in nozzle technology for water have resulted in laminar flow jets that give water the appearance (and optical properties) of a glass rod. So far, the optical properties of such water displays have not been exploited other than for their aesthetics. However, the specific optical properties of water jets, as well as droplets of water, lend themselves especially to image processing. In particular, the same arrangements of illumination, and laminarization of water, that give rise to the desirable appearance of the water, can also be used to concentrate illumination toward a camera system, so water can be easily tracked by image processing.

The goal of this paper is to describe the design and realization of simple water jet systems, based on image processing, that may be used as input devices for a variety of different tasks, such as typing or browsing information (such as web pages) in a wet environment. Of these, tasks, the most successful example was the use of water jets to create a new kind of musical instrument for use in an on-stage performance space, where typical stage lighting just so happens to have the effect of fully exploiting the special optical properties of laminar water jets and droplets of water that are ideal for image processing.

2. OPTICAL PROPERTIES OF WATER SPRAYS AND JETS

Fig 1 shows two differently illuminated pictures of essentially the same subject matter: a person running through some water sprays in a public splash area. The splash area consists of six hundred PEM Clearstream (TM) model number 824 nozzles, installed in the ground to shoot water upwards. These sprays were arranged, by artist Dan Euser, for aesthetic value, as well as for joggers and concert attendees to cool off in on hot summer days. The leftmost picture was taken, facing east, by observing the usual rule of photography, that the light should always be behind the photographer. All of the elements in the scene are properly exposed and visible.

The rightmost picture was taken facing west, with the late afternoon sun behind the subject matter. In this case, deliberately breaking the cardinal rule of photography, i.e. shooting into the sun, created a more dramatic shot in which the person was largely in silhouhette. Because the water spray acts like a lens, and bends the sun's rays, some rays of sunlight end up being refracted directly into the camera

---

Thanks to SPS energy, True North Power, Solar Roofing Systems, Diamond-Clad Power, and AO Smith, for help with this project.

0-7803-9134-9/05/$20.00 ©2005 IEEE

Fig. 1: A human interacting with water spray jets under frontal illumination (leftmost), and back-lit (rightmost). Backlit, jets and water droplets often behave like imperfect lenses that refract light over a sufficient range of angles to cause the water to be lighter than anything else in the scene.

by each drop or shaft of water, such that the water shows up much lighter than anything else in the scene. The shot on the right ended up being the one chosen for publication in various newspapers.

The purpose of this simple comparison is to show that the illumination that ended up being preferred for artistic reasons is also ideal for image processing of the water, since, under these lighting conditions (the picture on the right), the water ends up being lighter than anything else in the scene.

3. STAGE LIGHTING FOR A VERTICAL SPRAY WATER JET EXHIBIT/PERFORMANCE

A simple musical instrument for live stage performances was made from one jet of water, and one restrictometer, in which a performer plays different musical notes by blocking the jet with his or her hand, at various heights, to restrict the fluid flow in various ways. The stage setup is shown in Fig 2. The lights are arranged vertically because a laminar jet of water functions as a cylindrical lens that bends light primarily in azimuth but much less bending occurs in elevation. This lighting arrangement gave the best aesthetics for all of the audience members, since the audience members were seated approximately in a plane, i.e. varying widely in azimuthal viewpoint, but narrowly in elevational viewpoint. For safety, as well as portability, a small battery powered computer with a BT848 video capture card, connected to a 12 volt miniature NTSC video camera, was used for the real-time restrictometric image processing and instrument control. The entire system, including music synthesis and amplification runs from a 12 volt car battery, except for the stage lights which are 6 volt General Electric model 4515 lights.

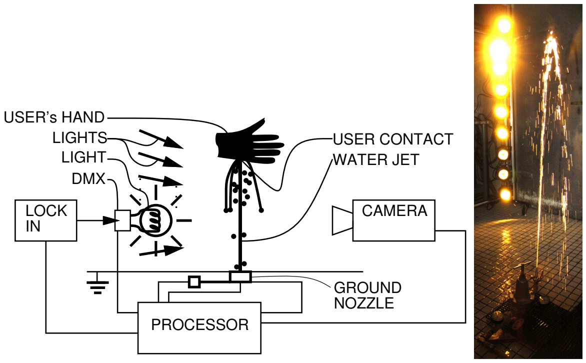

Fig. 2: Stage setup for water fountain performance system. The converging arrows each denote one of the nine stage lights arranged to point toward the audience as well as to point toward the camera that's used for real-time image processing. This arrangement serves to provide optimal aesthetics for the audience to view the water jet, as well as for optimal image processing. For simplicity, only one of the nine lights is shown in full. A lock-in amplifier was used to modulate the lights using the standard DMX512 protocol that is used for most stage lighting, theatre, lighting boards typically used at rock concerts, etc..

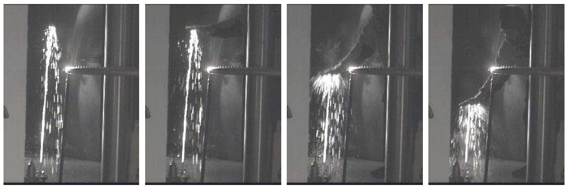

Fig. 3: Four frames of video from the playing of a typical song. Note the presence of three different musical instruments, in view of the camera, including a multijet instrument to the right, all being processed, in real time, from the one camera. The single jet musical instrument is visible in the left side of the video frame. Leftmost frame000 shows the jet with no hand present. This is the state for which no note is sounded. Next frame139 shows the hand just touching the top of the spray to sound the highest possible note. The rightmost two frames195 and 320 show the jet blocked at lower heights to provide lower notes. With image processing, it was thus possible to create a very sensitive restrictometer.

4. RESULTS

Four sample frames captured during a live performance, are shown in Fig 3. These are example snapshots from a real-time image processing system that computes the water column height (the time-varying strictometric measurement) as input to a musical instrument synthesis system, where the logarithmic frequency of the sound is proportional to the height of the water column.

Only a small portion of the image, of width 50 pixels, and height 300 pixels, is used for this single-jet musical instrument. The rest of the image is used for two other musical instruments, both of which can be seen to the right of the water jet, in the video frames. To the left of the water jet is a wall in the performance space that has the effect of blocking direct illumination from the stage lights, but the video camera has enough contrast discrimination capability that this barrier is not necessary to the functioning of the image processing system.

These narrow portions of the images, corresponding to the water jet, are first expanded in dyamic range. This is the first step in applying an anti-homomorphic filter [4][5]. Anti-homomorphic image processing works by first undoing the nonlinear response function, f, of the camera, (i.e.

expanding the already compressed dynamic range of the image, $f(q)$), to recover the quantity of light, q, then applying simple image processing to q, and then finally re-compressing (in dynamic range) the result.

The dynamic-range expanded narrow portions of the images are shown in Fig 4 together with the plotted results of a simple spray finding algorithm. The images are shown rotated, on their sides (jets pointing to the right), so that they will match the orientation of the corresponding plots. The algorithm to produce the data in these plots is very simple: first, an adaptive background subtraction is done based on statistical modeling of the environment. Background subtraction is well known. A good summary of various background subtraction methods may be found in Toyama et al [6]. Since the camera is fixed (stationary mounting), simple methods of background modeling, such as those presented in [7] are sufficient.

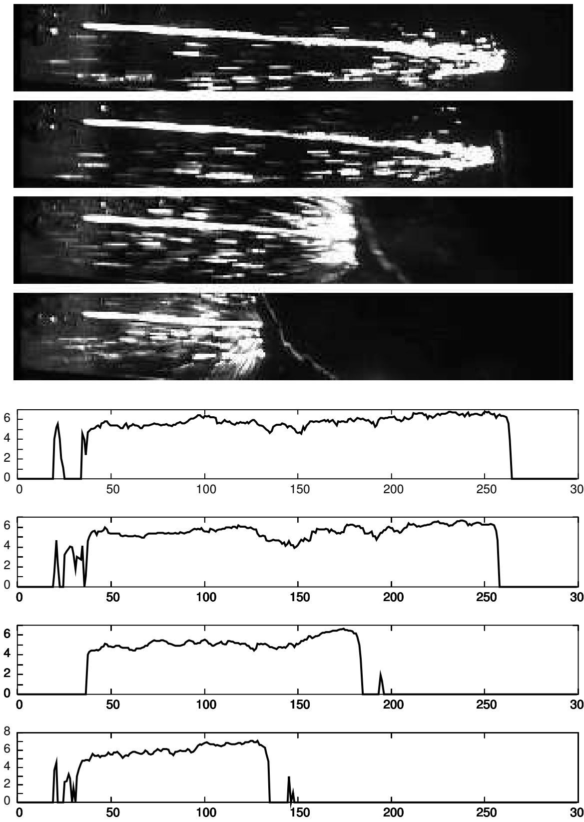

Fig. 4: Results of anti-homomorphic image processing on each of the four water jet images of Fig. 3. The relevant portion of each image is shown, on a dynamic-range expanded tonescale, at the top of this figure. These are displayed sideways to match the plots. The water column is clearly visible. The artifacts to the left are due to the nozzle and its feedback control system that is visible in the lower part of the image. In the lower two plots, the artifact to the right of where the water jet ends is the hand of the user. Note that the hand is more visible when playing low notes, because extending the hand further down shows more of the wet hand above the end of where the water jet column is blocked by the hand. The plots show, as a function of water column height (normalized to pixels), the logarithm of the pixel sum running across each jet, perpendicular to the direction of water travel.

Moreover, the first step of the anti-homomorphic image processing, namely the dynamic range expansion in going from $f(q)$ to q, tends to have the effect of emphasizing lighter areas of the image, i.e. the water sprays. This dynamic range expansion therefore functions much like a soft threshold, to subdue the effects of background areas, which, in our case, are darker than the water droplets and sprays.

Next a window is applied to emphasize the area of the image where the water sprays are typically found. This window has the effect of selecting the expected spray area from the whole image, which helps, for example, in selecting the one-jet instrument from the other water-based musical instruments in the performance space. This window also reduces the effect of background noise.

The combined effect of a window, $w(x, y)$ and the use of the dynamic-range expanded image $q(x, y)$ rather than $f(q(x, y))$ tends to localize in both the spatial domain (x, y) and the amplitude domain, q. Spatial windows [8] (localization in the domain of a function) and amplitude domain signal processing [4][5]. (localization in the range of a function) are both described elsewhere in the research literature.

This spatiotonal localization remains fixed because we know approximately where the water jet will be in the image, and we also know approximately the quantity of light that will be produced when and where the water jet is present.

The jet only sprays upwards. We wish to measure how far up, i.e. for what y value, there is sufficient quantity of light, $q(x, y)$ to consider there to be water at that height. This is done by summing horizontally along each row of the image after it has been weighted by the spatial localization window. Typically a truncated Gaussian window is used, although a Hamming, or Hanning window positioned and scaled to emphasize the fifty or so pixels over which the water jet is expected to be, tend to produce similar results.

Since the jet shoots upwards, the rows across the image essentially run perpendicular to the direction of the jet, so that, at each row, the windowing is one-dimensional. The jet is angled slightly, for aesthetic reasons, so the water curves up in a slight arc. This is not a problem because the adaptive background subtraction can still find it, and despite the small angle, horizontal rows across the image still cut across it at approximately right angles. The slight tilt of the jet just means that the window center needs to be adjusted so that, for example, in the images shown in Fig 3, the center of the window is about 25 pixels further to the left at the bottom of the jet, than it is at the top of the jet.

The weighted sum, along each row of the 300x225 pixel image:

$$g(y) = f\left(\sum_{x=0}^{224} w(x, y) q(x, y)\right) \qquad (1)$$

gives an estimate of how much water spray is present at each height, y. Notice the use of the function f which recompresses the dynamic range. A satisfactory function for f is the original camera response function, but more typically a logarithm was found to work equally well, to give:

$$g(y) = \log\left(\sum_{x=0}^{224} w(x,y)q(x,y)\right). \qquad (2)$$

This formulation also provides a meaning and interpretation more familiar to engineers than the seemingly arbitrary (though acceptably compressive) camera response function, f.

The first 40 pixels, which correspond to the nozzle jet, and various control system instrumentation (feedback loop to maintain constant jet height when the user is not interacting with the system, as well as provision to provide tactile feedback to the user) are truncated, so we only wish to look at g values from $y = 40$ to $y = 299$. These 260 values define the musical scale, over an approximately two octave range that is chosen to match the vocal range of the performer(s). This gives a resolution of approximately ten or eleven pixels per semitone on a standard equally tempered 12 note per octave scale. (In the situation where a particular song does not have such a full range of notes, the performer will often turn down the water flow, such as is shown in the figures, where the water jet does not go all the way to the top of the image area.)

It is evident from the plots in Fig 4 that we could easily just look for a region of contiguous values above a certain threshold, starting from $y = 40$, to get the value $y = 282$ for the top plot, $y = 275$ for the next plot, then $y = 197$, and finally, $y = 144$ for the bottom (last) plot. These values are robust, due to the effect of the anti-homomorphic processing, so they depend weakly on the value of threshold chosen. In fact any threshold value from $g(y) = 1$ to $g(y) = 4$ produces essentially the same results.

However, the negative derivative of g(y), more formal way of finding the point where the jet stream of water ends, was used:

$$h = \arg\max_{y}\left(-\frac{dg(y)}{dy}\right) \qquad (3)$$

This finds the value of height, h, as the value at which g drops off most quickly (i.e. the height variable y for which the quantity of spray ends most abruptly).

These values also agree approximately with the values above, e.g. for the bottom plot, $-dg(y)/dy$ has a maximum of 4.7 located at $y = 144$. Moreover, In order to get subpixel accuracy from the function $g(y)$, first moments of the negative derivative were used.

5. GOING FURTHER

The simple methodology of this paper was applied to a number of musical instruments and other devices that used fluid image processing for user input. For example, multijet implementations were made, in which each jet was assigned to a separate musical note, and amplitude, rather than frequency, was controlled by pressing down on each jet (Fig 5). These multijet instruments, together with the single jet version described in this paper were used in a number of successful performances in which the real time image processing was found to be robust and reliable.

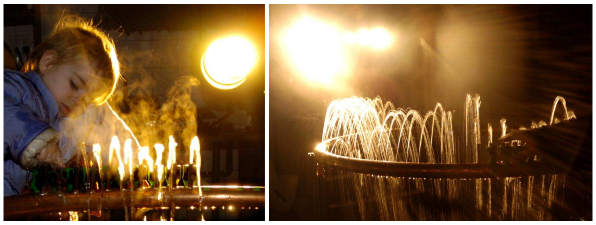

Fig. 5: Multijet versions of the new fluid-based musical instrument combine image processing with in-pipe restrictometers: a new application of image processing. The new musical instrument is simple to play. Leftmost: a 2 year old musician plays the well-known song "Jingle Bells" while the computer system responds robustly and accurately to detect, process, and generate each note in real time. Rightmost: a 61 note (61 jet) version of the instrument can be used to play over a 5 octave range with close jet spacing for easy chord formation.

6. CONCLUSIONS

The arrangement of standard stage lighting setups, for musicians and performers, when applied to the aesthetic display of water, created a situation that was particularly well suited to image processing. The result was a very simple image processing algorithm that was successful in tracking fluid as a user-interface medium, in real time, at 60 fields per second, on a small battery powered computer. This resulted in a safe and effective new input device suitable, for example, for use as a new musical instrument.

7. REFERENCES

[1] Meniere, "System for monitoring a swimming pool to prevent drowning accidents," in *U.S. Pat. No. 6,133,838*, October 2000.

[2] How-Lung Eng, Kar-Ann Toh, Alvin H. Kam, Junxian Wang, and Wei-Yun Yau, "An automatic drowning detection surveillance system for challenging outdoor pool environments," in *IEEE ICCV*, 2003.

[3] S. Mann, "Intelligent bathroom fixtures and systems," *Leonardo*, vol. 36, no. 3, June 2003.

[4] S. Mann, "Comparametric equations with practical applications in quantigraphic image processing," *IEEE Trans. Image Proc.*, vol. 9, no. 8, pp. 1389–1406, August 2000, ISSN 1057-7149.

[5] F. M. Candocia, "A least squares approach for the joint domain and range registration of images," *IEEE ICASSP*, vol. IV, pp. 3237–3240, May 13-17 2002, avail. at http://iul.eng.fiu.edu/candocia/Publications/Publications.htm.

[6] Kentaro Toyama, John Krumm, Barry Brumitt, and Brian Meyers, "Wallflower: Principles and practice of background maintenance.," in *ICCV*, 1999, pp. 255–261.

[7] Christopher Wren, Ali Azarbayejani, Trevor Darrell, and Alex Pentland, "Pfinder: Real-time tracking of the human body," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, July 1997, vol. 19, pp. 780–785.

[8] F.J. Harris, "On the use of windows for harmonic analysis with the Discrete Fourier Transform," *Proc. IEEE*, vol. 66, no. 1, pp. 51–83, Jan. 1978.