Pressing the Flesh: Sensing Multiple Touch and Finger Pressure on Arbitrary Surfaces

Joe Marshall, Tony Pridmore, Mike Pound, Steve Benford, and Boriana Koleva

Mixed Reality Lab, University of Nottingham, NG8 1BB, UK
{jqm,tpp,mpp,sdb,bnk}@cs.nott.ac.uk

Abstract. This paper identifies a new physical correlate of finger pressure that can be detected and measured visually in a wide variety of situations. When a human finger is pressed onto a hard object the flesh is compressed between two rigid surfaces: the surface of the target object and the fingernail. This forces blood out of the vessels in the fingertip, changing its colour slightly, but systematically. The effect is visible to the naked eye and can be measured using techniques from computer vision. As measurements are made of properties of the hand, and not the target surface, multiple-touch and pressure sensing can be added to a range of surfaces - including opaque, transparent, smooth, textured and non-planar examples - without modification of the underlying physical object. The proposed approach allows touch sensing to be fitted to surfaces unsuitable for previous technologies, and objects which cannot be altered, without forfeiting the extra range of expression of pressure sensitivity. The methods involved are simple to set up and low cost, requiring only a domestic-quality camera and a typical computer in order to augment a surface. Two systems which exploit this cue to generate a response to pressure are presented, along with a case study of an interactive art installation contructed using the resulting technology. Initial experiments are reported which suggest that visual monitoring of finger colour will support recogntion of push events.

1 Introduction

Touch sensitive surfaces, such as graphics tablets, interactive whiteboards, touch screens etc. have existed for some time. Touch sensitivity, however, typically requires the surface to be enhanced with some kind of embedded electronics, or in the case of capacitive sensing on glass [1], to have electronics below the surface. Computer vision has the potential to create touch interfaces without embedding electronics in the target surface, and also to detect multiple touches. Current systems, however, are typically either unable to detect the difference between touching and moving a hand or object near the surface [2], or can only detect the presence of a finger or object next to the surface (by using cameras at right angles to the surface [3], or multiple cameras and some form of 3d disparity measurement [4,5]).

The main contribution of this paper is to identify a new physical correlate of finger pressure that can be detected and measured visually in a wide variety of situations. When a human finger is pressed onto a hard object the flesh is compressed between two rigid surfaces: the surface of the target object and the fingernail. This forces blood out of the vessels in the fingertip, changing its colour slightly, but systematically. Increased pressure increases the effect, up to a limit determined by the thickness of flesh on the finger. Colour change may be seen either by examining the pattern of colours in the fingernail or, if the target surface is transparent, by looking at the fingertip through the surface. When viewed through a transparent target surface, increasing pressure increases the amount of flesh from which blood is expelled, creating a larger region of paler skin. When viewed from above the hand, through the nail, increased pressure forces more blood to the base of the nail, concentrating colour there. Both these events are clearly visible to the naked eye.

In what follows we describe computer vision-based sensing methods which exploit this cue. As the approach relies on measurements of the physical properties of the hand, and not the target surface, it has the potential to add multiple-touch and pressure sensing to a range of surfaces - including opaque, transparent, smooth, textured and non-planar examples - without modification of the underlying physical object. This allows for many new items, such as stone carvings or wood, to become touch sensitive interfaces. No technology need be embedded into the target object; all that is required is that a colour camera be positioned to view the effect. The proposed method is therefore potentially highly flexible, easy to install, low cost and portable. It requires only a domestic-quality camera and a typical computer in order to augment a surface.

The proposed approach is expected to be of particular use in environments such as museums, science centres and galleries. Here, visual sensing of finger colour can allow people to interact directly with existing physical objects, or with glass cabinets containing objects of interest, without having to customise the objects or cabinets themselves. The method has benefits for installation designers, allowing museum and science centre staff to construct interactive exhibits and environments based upon their existing catalogue of objects. For example, historic tools in a countryside museum could be touched in order to trigger audiovisual material about their use. The flexibility of the approach means that exhibits could also be reconfigured easily and on a regular basis, maintaining visitor interest.

The paper is organised as follows. Section 2 briefly reviews relevant prior work, before Section 3 describes finger pressure sensing by viewing the tip of the fingernail from above. Section 4 then describes a method which uses the same visual cue, but views the hand from the rear of a transparent glass surface. The fingernail-based method was used to create an interactive art installation, in which the user was able to interact with a pressure and touch sensitive rock. This installation is presented as a case study in Section 5. A key motivation for the development of touch sensitive interfaces is the ability to detect touch events such as contact, pushes, taps, etc. Initial experiments are reported which suggest that visual monitoring of finger colour will support recogntion of push events are described in Section 6. Finally, conclusions are drawn in Section 7.

2 Prior Work

Touch sensitive graphics tablets [6] and touch screens [7] are widely available pressure sensing interfaces. The most common pressure sensitive interface in production is the laptop touch pad. Whilst these are typically used purely for on/off touch pressure, most, such as those made by Synaptics [8] also are able to detect variations in pressure. Some models, for example the Mitsubishi DiamondTouch table [9], even allow the detection of multiple touches, although not pressure.

Following a different approach, Schmidt et al. [10] used load sensors to create a touch sensitive table. In addition to touch events, they detect several contextual events such as objects being put down on a table, which is interesting as it relates to our goal of augmenting existing objects. Schmidt et al, however, cannot support interactions with objects other than moving them around on the table. This does not require additional technology in the sensed object itself, but load cells are required to be fitted at four corners of the surface. It is also limited to single touch interaction on horizontal surfaces.

Exploiting vision and related technology to create a touch detecting screen is not a new idea. Various methods have been used, such as scanning laser rangefinders [11], internal reflection inside a glass plane[12], multiple cameras and planar homographies to detect only pixels that are near the screen [5], and the visual detection of (somewhat exaggerated) finger gestures in order to detect touches on a virtual keypad [13]. These visual methods typically fail to detect pressure differences during touching, although some level of pressure sensing has been demonstrated with the internal reflection method, by using the size of the finger's contact area. The finger surface area is also used in Benko et al's multi touch table[14]. Benko et al suggest it is too innacurate to detect pushing reliably and define a special rocking gesture for clicking which their system is able to detect. The two sided LucidTouch system[15] also uses visual tracking to detect the hand position, however it uses a separate touch sensitive pad in order to to detect touches on the surface (as the vision tracking method used is unable to detect touch).

These technologies are designed for use in interactive whiteboard, wall display or table interfaces. They usually require modification of the sensing surfaces in some way, or place restrictions on the surface being monitored. They are also currently designed for completely flat user interfaces. This may be suitable when used as an interface to standard GUI style applications; however as interface designers move beyond the GUI, this may become a limitation. When augmenting existing objects, it is hard to guarantee complete flatness. Bumpy or angled surfaces may also be useful to allow tactile feedback as to where the hands are, which is commonly seen as a reason why touchscreen interfaces such as virtual keyboards have only had success in niche applications.

3 Fingernail Sensing

When the fingertip is pressed down on a surface, the blood under the nail concentrates at the bottom of the nail, and the tip of the nail becomes whiter (Fig. 1). This effect is very consistent, and only requires a small amount of finger pressure for a difference to be clearly visible to a human observer. This section discusses the automatic visual detection of this cue.

Fig. 1. Nail at different pressures

The fingernail sensing system uses a basic background segmentation algorithm, followed by a contour detection operation to find the fingertips. When a finger is detected which has not moved more than a small threshold since the last frame, the image of the fingertip is examined, and the distribution of colour in the nail quantified. This reflects the pressure exerted by that finger.

Initial attempts at sensing pressure used the two parts of the nail, the tip and the bottom, and compared the colours of these to detect a change. However, the exact location of the white areas on the tip of the finger proved to vary significantly between individuals, and is also difficult to sense from any distance. For example, when viewed from 60cm with a 320x240 pixel camera, the nail is approximately 10x12 pixels in size, which means that the tip area in particular is too small. However, while the fingernail is almost uniformly coloured when no pressure is applied, two distinct colours appear on the nail when pressure is exerted. Because of this, rather than use located features on the nail, we simply take the variance of the hue of the pixels in the nail area. In order to calculate a mean hue, the hue is represented as a 2d vector, and an arctangent applied to this. Variance is calculated with an allowance for the circular nature of the hue metric.

$$MeanHue = \text{atan2}([\sum_{1}^{n} \cos(Hue)], [\sum_{1}^{n} \sin(Hue)])$$

$$VarHue = \frac{1}{n}\sum_{1}^{n} min((Hue - MeanHue)^2, (360 - (Hue - MeanHue))^2)$$

Initial testing has shown this metric to relate strongly to finger pressure. It is also much more detectable at a distance, and produces similar results on different fingers. Variance of the brightness of the pixels can also produce useful data in some conditions, however it is, as might be expected, extremely sensitive to illumination changes. When the finger is pressed down hue variance clearly increases, with the opposite effect visible on release. Pushing less hard produces an intermediate response. Because the blood under the skin moves back into its normal place relatively slowly, there is a natural smoothing on the release of approx 100ms, this may be useful for 'debouncing' purposes, avoiding multiple presses being detected when the finger is only pushed down once.

It is also clear that depending on lighting and individual variation, the absolute variance values alter somewhat. A floating normalisation window is therefore employed, with the value of 'pressure' detected being mapped to 0...256, by using previously recorded pressure values as a max and min. A constant minimum pressure range (mr) is used, for the case when the finger is first seen, and only a small amount of data is in the window. This avoids large random fluctuations if the finger is simply placed down and not pressed. The raw to normalised conversion is expressed as:

$$normalised_t = \frac{256 * (raw_t - \left(\min\limits_{k=t-windowSize}^{k=t} (raw_k) \right)}{\max \left[mr, \left(\max\limits_{k=t-windowSize}^{k=t} (raw_k) - \min\limits_{k=t-windowSize}^{k=t-1} (raw_k) \right) \right]}$$

This assumes that when the finger is first seen, there is no pressure on it, which is the case in typical use; even if a press is occurring the finger is first seen as the press starts. This conversion, whilst it means that no exact pressure data is available, makes push and release events clearly visible, and allows intermediate pressure values to be acquired. Normalisation is effective as long as the raw variance is altering with finger pressure. It has a compressing effect on the raw curves, which was desirable in our application (Section 5), but may or may not be suitable depending on context.

3.1 Initial Evaluation

In order to test the fingernail algorithm, a test rig was constructed, with the user's finger pushing on an electronic scale which served as a ground truth pressure gauge. The output from the scale was then video recorded along with the output from several brief sessions of pushing and releasing a single finger. The test rig was able to detect a 'weight' of 2kg (a force of approximately 19.6 Newton). In practice, this limit was not a problem, as forces outside this range proved uncomfortable to apply. The scale reported weight with a relatively slow update rate, updating at up to 4 times a second. Measurements from the visual system were taken each time the scale's reported weight changed. The system was run, and the hand moved into view until the hand tracking found the finger, and then the output was recorded for approx 50 seconds.

Single User Reliability Once the data had been recorded, the raw hue variance data for each test session was scaled in order to make the mean and standard deviations the same as the ground truth. These normalised graphs showed a very good fit to the ground truth data, with a certain amount of clipping at the highest pressures in some tests. These results were analysed using regression analysis, which gave a P value of <0.1% for all subjects. Figure 2 shows 3 different user's normalised pressure outputs plotted against ground truth. These graphs demonstrate the ability of the algorithm to reflect several push release cycles accurately. Detecting an initial push is possible, as the minimum pressure range means that an initial push will have a different profile to just touching (the first graph in figure 3 starts by just touching the surface, whereas the second and third graphs start with a push, the visual measure ramps up high straight away. The tracking works at 30 frames per second, limited by the camera frame rate, rather than any processing constraint, which is fast enough to detect quick push and release cycles.

Between User Variation and Lighting Variation Tests were carried out on different days, in a naturally lit room. This meant that the system was exposed to some lighting variation. To quantify the possible effects of lighting variation, one user was tested on two different days, both times using the same finger.

The variation in lighting had a major effect on the raw variance values from the system, with the same user showing a significantly lower range of variances, which were also significantly higher than their previously recorded values. Multiple users in the same lighting conditions also had differences in the distributions of variance, although these were significantly less than the lighting induced variances. Figure 3 shows some examples of these effects.

These two factors mean that unless very controlled lighting is available, and a training session is undergone for each user, this method is not suitable for providing absolute pressure information, ie. it is not suitable to replace a load sensor. However, when normalised as described above, it can be employed in interfaces where a correlate of pressure, rather than true pressure, is required. It seems likely that colour-based measures can support detection of more fuzzy actions such as pressing, pressing hard, pressing softly etc., as is required in most touch based interfaces.

When Does This Work? Several factors may cause this method to fail. Firstly, nail varnish or gloves will obviously cause the system to fail, as the fingernail cannot be seen. Secondly, if the fingernail is very brightly lit by direct sunlight, this may reflect off the nail, making it impossible to see the skin colour beneath it. This was the case during one of the test sessions, with the system failing to work until a curtain was drawn to block the bright rays of sunlight.

The method is reasonably robust to changes in finger orientation. As long as the length of the fingertip can be seen, a correlate of pressure is produced and changes in the hue varaince reflect changes in pressure. If the angle is changed during sensing however, the values can be seen to change slightly. This means

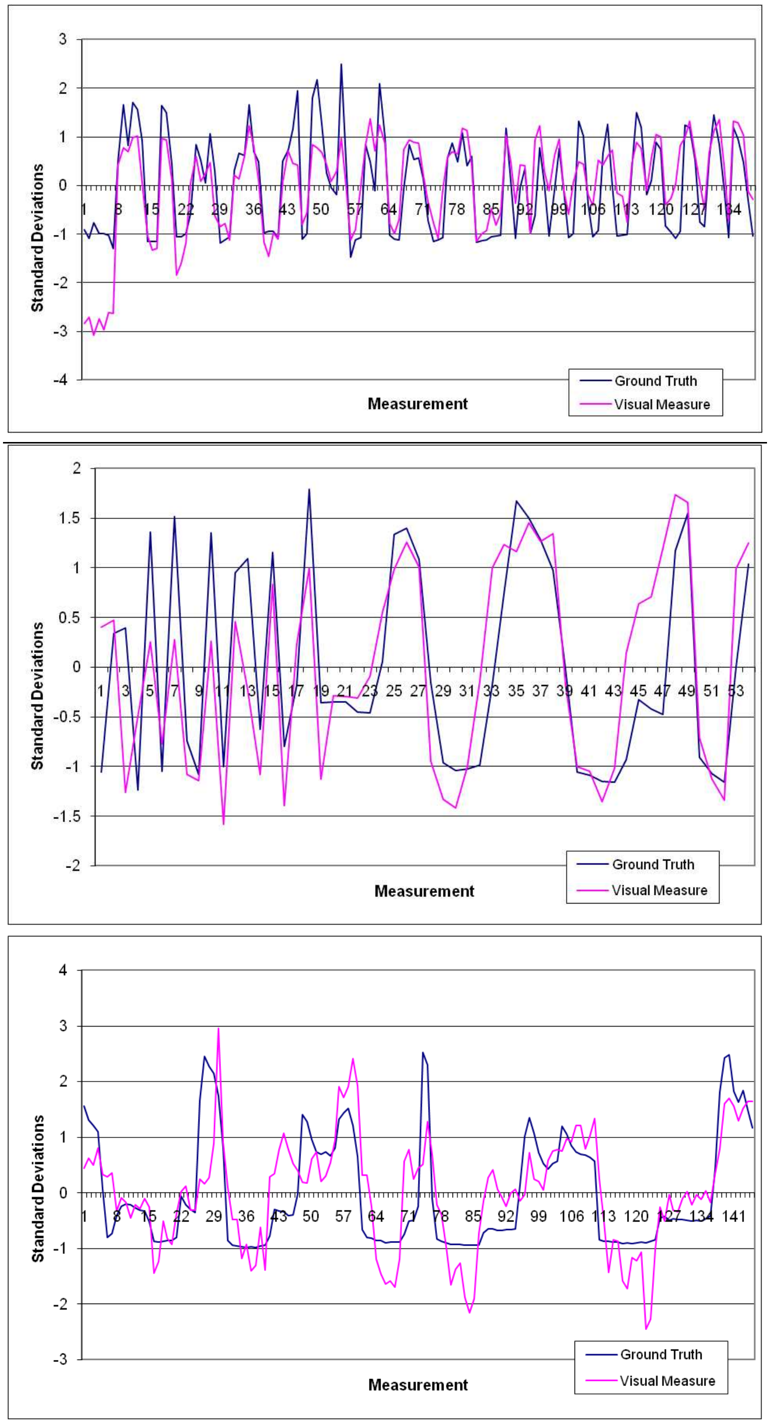

Fig. 2. Examples of Performance of Pressure Tracker over Time

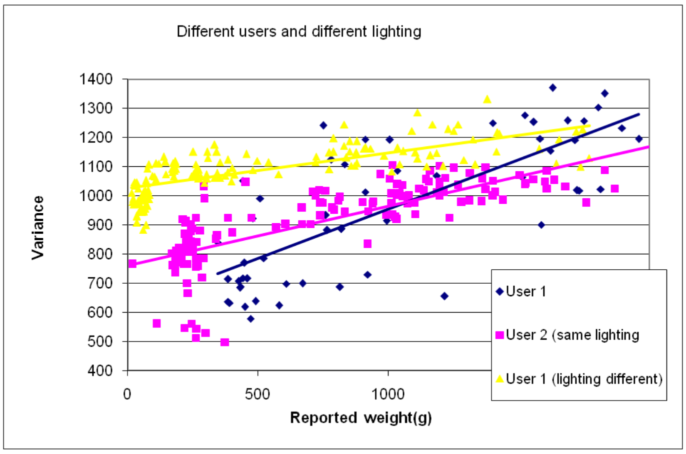

Fig. 3. The Effect of Users and Lighting on Fingertip Variance

that there is potential for use on non-flat surfaces, as long as the finger is not changing in angle massively during a single touch movement. A slight side effect of our simple hand tracking system is that when the fingers are clasped round an object, so the fingernails are out of view, the knuckles and what is visible of the finger above them are detected to be fingertips by the system. When the knuckles are detected as fingertips, the system still responds, as the knuckles are differently coloured to the rest of the finger, and grasping causes the ratio between the knuckles and the part of the finger that is visible to change, thus altering the variance of the detected 'fingertip' (see Figure 4) . Potentially useful data is also provided if the hand is held in the air, and the thumb is squeezed against the bottom of a finger.

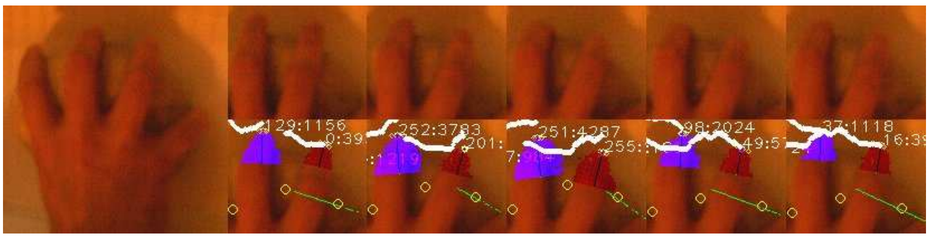

Fig. 4. Grasping and releasing an object. Each frame shows the zoomed in middle finger, and the pressure graphs next to the middle and index fingers

4 Skin on Glass

To assess the potential of visual monitoring of skin colour to detect pressure on transparent surfaces, such as windows, glass cabinets, etc., the same approach was applied from the other direction, tracking the finger through a sheet of glass. Changes in skin colour were recorded as the finger was pressed against the glass. It was found that at the point of contact pressure was sensed reliably if normalised as described in Section 3. Intermediate pressures were again detectable, and pushing and not pushing generated distinct output profiles. Sensing through glass may be of particular value as it allows the computer and camera to be entirely enclosed, for example behind a shop window, or inside a glass case, with no exposed electronic parts. It also has an advantage over the fingernail tracking in that it is less susceptible to occlusion, which may be a problem in some uses of the fingernail method.

This technique is, however, not quite as reliable as the fingernail-based method in one particular: until the hand is touching the glass, sensing is somewhat erratic. Further research is required, but this is probably the result of the changing distance between the glass surface and user's hand. It seems likely that the fingernail-based method is more reliable because the nail is firmly attached to the surface of the finger, so that the relationship between the fingertip and the surface through which it is viewed remains constant. In the test application, the effect is reduced by only starting to record pressure once the detected fingertip has been in the same position for 3 frames. This means that there is a delay of approx 1/10th of a second before continuous pressure readings begin. It also means that if the hand is held very still in the air in front of an interface, pressure sensing will begin, although it will only break if the hand is very slowly moved directly towards the camera, which proved hard to do in testing. An output from this sensing during two pushes on a rather dirty and reflective sheet of glass is shown in Fig. 5 (this version was tested with a black background, as skin segmentation did not prove a problem in the initial skin on glass tests). The method works well even in sub-optimal conditions. Figure 5 also shows a graph showing the comparison of the skin on glass to the ground truth measurement (as used in Figure 2 for evaluation of the fingernail tracker.) Note that this system uses an identical algorithm to the fingernail tracking, with the normalisation taking care of the smaller absolute variance values seen in this method. It is possible for a user to simply move their hand to the other side of the glass and use the fingernail tracking without any recalibration.

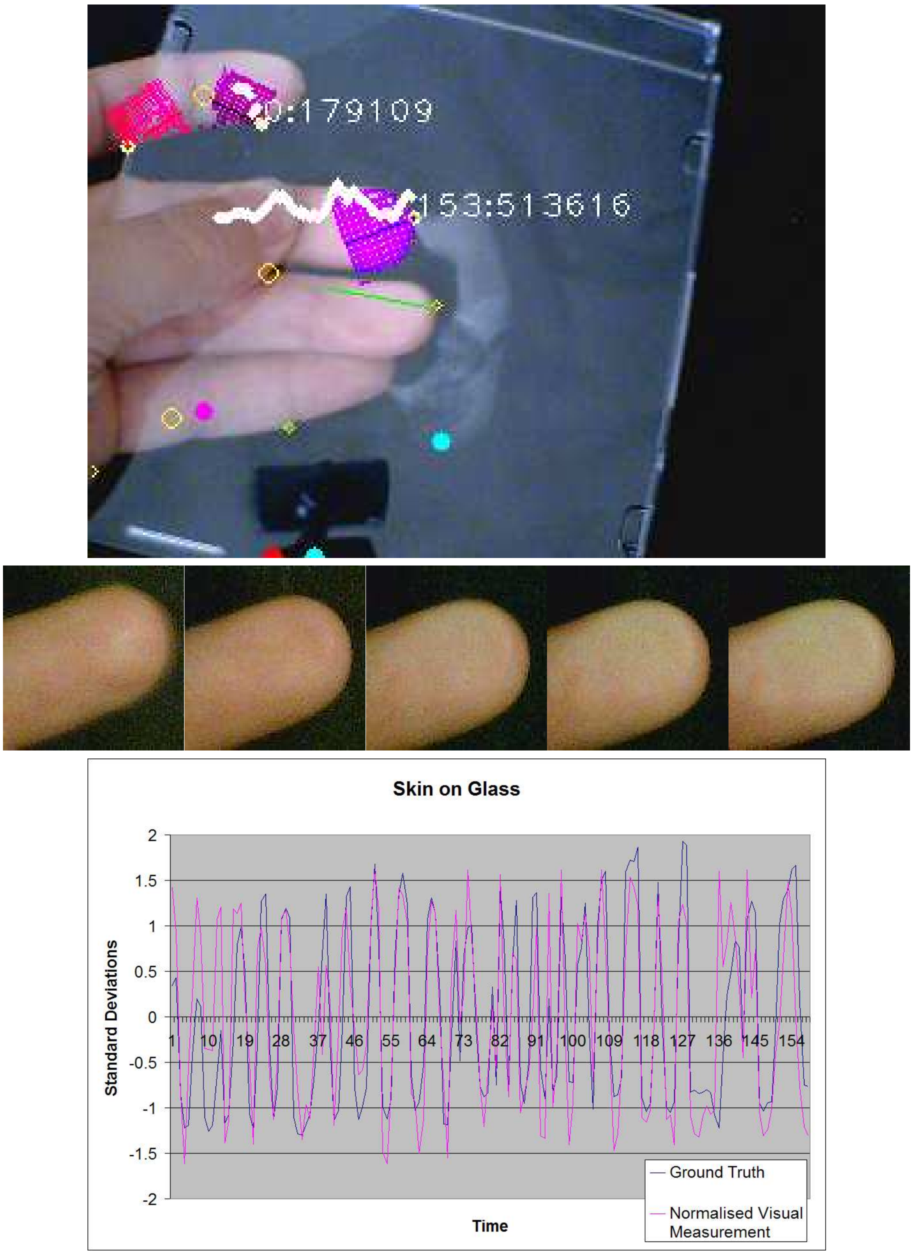

Fig. 5. Sensing pressure through glass - the line of pictures shows the fingertip over a single push sequence.

5 Case Study: Rock

The fingernail tracking algorithm was evaluated further in an interactive art installation called Rock. The installation presents a rock in a cage as a pet. A web camera is attached to the cage's top, and a computer and speakers are hidden under a table that the cage is on (see Fig. 6). Rock uses gestural input and audio output to mimic the personality of a small pet rodent such as a guinea pig. It is designed to have quite a timid personality and to be easily frightened. The rock is an extreme test of the behaviour of the fingernail algorithm with a large range of gestures and angles, and provides a testing ground for graceful fallback in situations where it is impossible to sense pressure.

Initially the rock makes a quiet steady heartbeat sound. When the rock is touched it responds by making animal sounds, and the heartbeat changes to signify its level of fear. Touching the rock in different ways can provoke varying responses, for example if touched gently and slowly, it is likely to make quiet purring noises and not be very scared: grabbing at the rock too quickly scares it and makes it snarl or growl.

The rock is designed as an ambient installation, to be left in a gallery or space at an event, and interacted with by people with a minimum of direction. As such, it is designed to attract people to interact with it; this takes two forms. Firstly the heartbeat sound attracts interest to the rock when it is not being interacted with. Secondly, the interaction with the rock is designed to be interesting to onlookers. The interaction is designed so that onlookers can see part of the way that the rock is being interacted with, but so that part of the interaction is not visible to them. In particular, the finger pressure detection is used here and provides an aspect of the interaction that is unclear to onlookers, and designed to intrigue people into interacting with the rock themselves.

Fig. 6. The Physical Setup of the Rock

5.1 Technology implementation

The single camera on the top of the cage is the only input mechanism for the rock. The camera is carefully positioned so that it can see all of the bottom of the cage. All the output comes from the speakers, which are positioned so that the sound seems to come from the bottom of the cage.

The computer detects the silhouette of a hand reaching into the cage by use of a simple colour threshold to select pixels which match the colour of the background or the rock. There is deliberately no skin colour detection, or scene based background subtraction, in order to make the system responsive to non-skin objects put into the cage (as long as they are not the same colour as the background or the rock), and also to allow the rock to be moved in the cage without breaking the background model. This makes for a very reliable and simple detection of the hand silhouette when the hand is inside the cage.

The system detects how close the hand is to the rock, and uses this over time to calculate a measure of how fast the person's hand is approaching the rock when they reach into the cage. It also attempts to find the fingertips and detect the average finger pressure on the rock over all the fingertips it can see, by using the algorithm described in Section 3.

These two measures, of approach speed, and pressure are mapped respectively into two variables, 'fear', and 'excitement'. These variables are mapped onto a set of audio samples, and audio processing filters which alter these sounds. The audio samples used were made by one of the authors, and are categorised as to how scared and how excited they sound. The audio processing effects are a mixture of time and pitch shifting, and are used in order to make the sounds sound different every time they are played rather than like a fixed set of samples. Examples of sounds that the rock may make are a low growl if it is scared but not very excited, a high pitched snarl if it is scared and excited, purring sounds if it is not afraid but not very excited and squealing sounds if it is excited and not afraid. A slight element of randomness is added into this mapping; this is designed to make the rock be mostly predictable, but to avoid letting the users be certain how it will respond to a particular gesture. The heartbeat sound continues all the time beneath the animal sounds, getting faster and louder when the rock is more scared.

5.2 Testing the Rock

The rock was exhibited at a recent digital art conference. At this event, it was placed in a corner of a corridor space, where a lot of people were passing by, as an ambient installation during the conference. This allowed us to see the rock interact with approximately 100 people, from various backgrounds including art, architecture, sound design, HCI etc. The rock was running for over 9 hours, and was very successful in this environment; the installation was awarded best paper prize.

Initially, one of the authors was with the rock, introducing it as his pet, in order to entice people to play with it. After a few people had played with the rock, this became unnecessary, as people started bringing back other people to show it. At this point, the rock became more interesting, as the explanations people were creating for its behaviour became increasingly complex and rich. At the end of the event, one of the participants was very attached to the rock and even asked if she could take the rock home. The descriptions of the rock and its personality were very varied, ranging from ‘cute’ to ‘strange’ and ‘disturbing’.

There were several ways in which people interacted with the rock. Most common initial interactions were poking it, either suddenly, or gingerly reaching in to touch it. In these cases, the technology responded reliably. Once people had realised that the rock was not going to bite them, they explored more complex interactions, such as stroking it (which worked as long as they didn’t move their hands too fast), and grasping it. Grasping was interesting, because the effect discussed in Section 3.1 meant that the knuckles were tracked, giving a pressure signal as to how hard the rock was grasped. This meant that in this (relatively common) mode of interaction with the rock, the tracking still worked, although slightly less reliably. A few people did things such as picking up the rock from underneath, waving their hands right in front of the camera, or closing their hand into a fist when touching the rock. In these situations, the pressure tracking broke, and the rock responded to the movements using only the silhouette of the visible part of the arm, which led to slightly unpredictable responses to these particular movements; it was important in the design of the rock’s ‘personality’ that it handled the cases when finger tracking data became unavailable, and still provided some kind of response. The unreliability when presented with these odd gestures was translated in the user’s eyes to become a facet of the Rock’s personality, for example as it not liking having strange things done to it.

While observing the rock, it was clear that the balance between the unpredictable nature of its response to odd actions, and the predictable response to actions such as stroking gently and holding it, formed a part of the success of the installation. The ambiguity allowed people to spot ‘patterns’ and create explanations, and meant that whilst people could to some extent learn things about how to control the rock, such as not to grab at it and scare it, or by using gentle

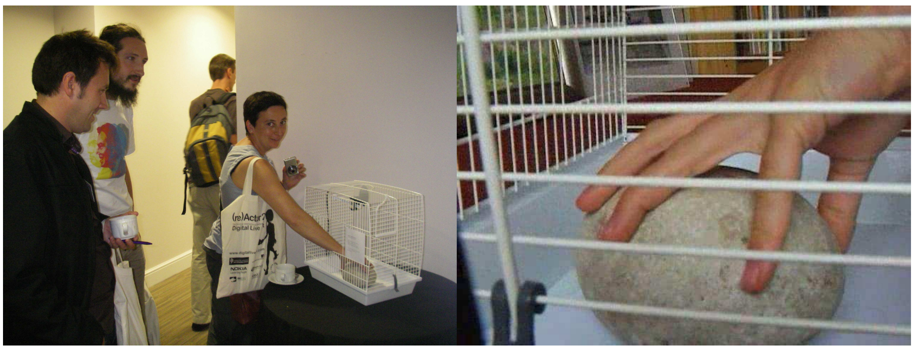

Fig. 7. Interacting with the Rock

touches to make it happy, they were not able to get to a level where they felt they had complete control. One important thing however is that the level of reliability was such that the 'owner' of the rock was able to demonstrate that the rock 'liked' him, and that people were able to learn how to touch it to make it likely to make 'happy' sounds.

The finger pressure sensing method was important in this installation, as it allowed a very expressive mode of interaction with the rock, but without having to augment the rock with sensors. Within the constraints of the cage, this created effectively a wireless, remotely powered, touch sensitive moveable user interface, which was made of seamless stone, with no charging connectors or battery compartments. Alternative ways to create similar effects would have created points at which the audience's suspension of disbelief was broken. For example, a pressure sensor under the rock or cage would fail to work if the rock was lifted, adding sensors to the rock itself would be hard to do without external electronics, battery compartments etc. which would break the concept of it being an organic creature.

6 Using a Bayesian Classifier to Detect Push Events

A key motivation for the development of touch sensitive interfaces is the ability to detect touch events such as contact, pushes, taps and double clicks. To provide an initial indication of the feasibility of detecting such events given colour variance data a bayesian classifier was implemented and used to detect contact between hand and surface. Colour-based detection and location of human skin is now commonplace in computer vision systems. A number of skin detection techniques have been reported [16,17,18,19], most based upon the work of McKenna et al. [16] which showed that colour spaces exist in which, for a wide range of nationalities and ethnic backgrounds, human skin is tightly clustered. The classifier to detect contact between human fingers and a target surface by identifying compressed flesh adopted a similar approach.

A camera was placed behind a sheet of non-reflective, smear-resistant glass, providing a clear view of the user's hand as s/he interacted with the other side of the surface. The challenge was to use the resulting colour images to recognise the differences between:

- the normal, i.e. uncompressed, skin seen when the user's hand is in view, but not in contact with the glass;
- the compressed flesh that appears when the user touches the glass surface;
- the environment behind the user.

All experiments were carried out in an office/laboratory environment, so the background comprised arbitrary coloured objects. Some of these objects were approximately skin-coloured, but no other people (i.e. no additional real skin) was allowed into the field of view. Six individuals, of mixed age, sex and race took part. Each was first asked to press his/her hand flat onto the glass panel to provide easily identifiable examples of contact, and then invited to press, tap, or otherwise touch the glass at will. Two minutes video of each subject was captured and analysed off-line.

Following [16,18], the well-known hue-saturation-intensity (HSI) colour space was employed throughout. The hue (H) and saturation (S) values associated with human skin are known to cluster tightly, though intensity (I) varies widely. Bayesian classification was used to separate the three classes (uncompressed skin, compressed skin, non-skin) identified above as shown in Fig 8. Models, in the form of approximations to probability density functions for each class were first constructed from manually identified training data. This classification of pixels into tip and non-tip could potentially allow for reliable detection of touch pressure, by detecting the size of the compressed region of the fingertip. Raw colour values were examined to determine whether or not the information required was present, before any features or summary statistics were computed, in the base image data.

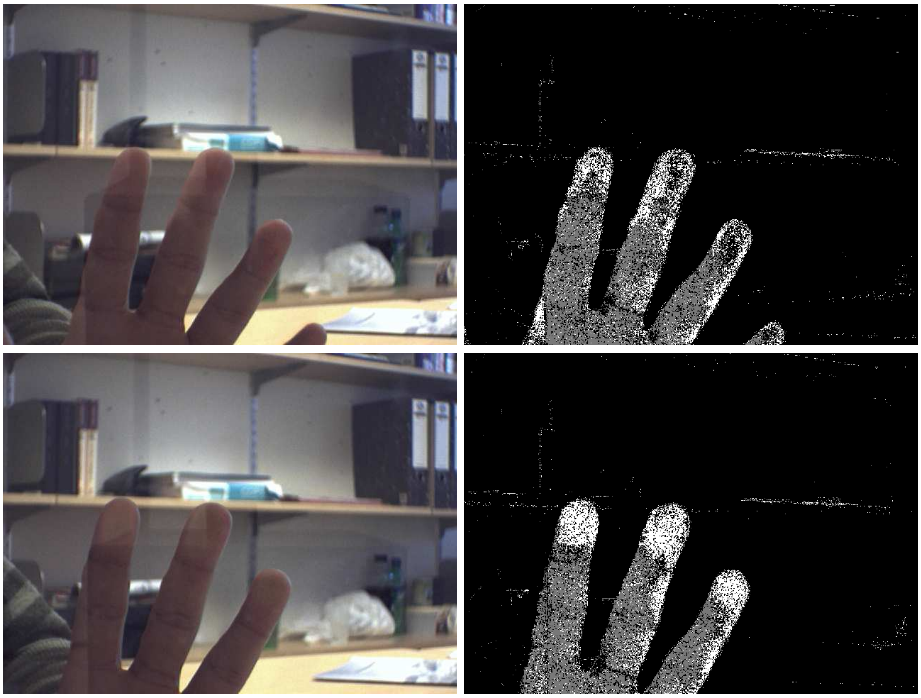

Fig. 8. Skin on glass - hand just touching glass (1), and fingers pushed against (2) - grey pixels are classified as non compressed skin, white pixels are compressed skin.

When trained on data from an individual's hand, bayesian classification was found to be effective. Contact between fingertip and glass could be reliably detected. However when applying the same algorithm to multiple users, by pooling training data to produce composite colour models, it was found that the variation between users was often equal to the difference between compressed and non-compressed skin for a single user. The experiment therefore demonstrated that variations in individuals' skin colour can better support automatic detection of touch events. As a result, further work on event classifiers will exploit time-based measurements of individuals' finger pressure, similar to those we have used for smooth pressure sensing.

7 Conclusions

We have demonstrated a new correlate of finger pressure which can be measured visually using standard equipment in a wide variety of circumstances. The method detects compression of the fingertip by monitoring changes in the colour of either the skin or the fingernail. Table 1 summarises the main strengths and weaknesses of this approach to pressure sensing.

Strengths	**Weaknesses**
No modification of tracking surface required	Viewpoint and occlusion
Quick, easy and cheap setup	Lighting
Smooth pressure sensing	Relies on hand tracking
Potential to support automatic detection of touch events	Not fully 3D
Wide range of surfaces can be augmented	
Multiple touch	

Table 1. Benefits and Challenges of using Fingertip sensing

The method allows the addition of an extra dimension of expressiveness to previous vision based hand & finger sensing systems, without requiring complex addons such as multiple cameras, or augmenting the surface in any way. It is inherently multiple touch, as it measures a feature of the pressure on the finger, rather than the pressure on the surface below the finger. It is quick, easy and cheap to setup. The technique extends the range of materials and surfaces available to standard pressure sensing, by allowing touch pads to be created from any relatively firm surface which is visible to a camera. As demonstrated in the Rock example, fingertip pressure sensing does not even require a flat surface, working well when given a bumpy surface. Though further development is required to produce a working system, the colour measures employed here clearly have the potential to support detection of a variety of touch events.

The approach is, like many vision-based techniques, potentially sensitive to camera viewpoint and occlusion and is unlikely to work well in some extreme lighting conditions (very bright sunlight & darkness). It is also reliant on the hand tracking working correctly in order to function; if the tracking fails, no pressure sensing can occur. It is not fully 3D, as it cannot sense pressure when the whole of the fingers are out of view, so applications have to be designed to degrade gracefully if this is a possibility, however it provides useful data in a large range of situations, such as when the fingertips themselves are out of view, as long as the grasping hand and the rest of the fingers are still visible to the system. Also, whilst it works on a wider range of surfaces than most current systems, there clearly is a limit to what surfaces it can work reliably on, for example surfaces such as cushions, gels or liquids will all be impossible to augment.

7.1 Potential Applications

Whilst this technology clearly may be useful in tabletop displays and other common multi-touch interfaces, it has most to offer in the creative, museum and educational sectors. The ability to augment an existing, everyday, physical object would be of particular use to museum, science centres, exploratoria and other similar places where a hands-on approach is encouraged. The skin on glass method of touch sensing provides a useful extra mechanism for objects which are in cabinets and unable to be directly touched. In this situation the hardware would be fully enclosed within the cabinet, which may be an advantage. Augmenting unexpected surfaces in this way has proven interesting and surprising to users in our case study; it is envisaged that in a museum setting, being able to augment the object rather than having a separate interactive display may provide a more direct and engaging experience.

As well as being useful for currently impractical applications, the techniques reported here make pressure and touch sensing available with a significantly lower setup time than existing methods and require no custom equipment; the Rock takes approximately 5 minutes to install and uses a cheap domestic webcam and PC. This means that the proposed method has the potential to be incorporated in mass market entertainment software, for example this could enable innovative interfaces such as used on the Nintendo DS touch screen game console to be created on a larger scale for home users (For example in Warioware Touched, users have to 'rub out' on-screen pictures, stroke dogs, whack moles etc. by using touch gestures).

7.2 Future Work

The work described here has demonstrated the potential of visual monitoring of skin colour to reflect finger pressure in a range of situations. Topics for future research include:

- investigation of alternative methods of capturing changes in skin colour, and their relation to finger pressure. In particular, though the current method is reasonably robust to changes in finger orientation it is not invariant under such changes.
- evaluation of the usability of the approach in a wider variety of application domains and scenarios, focusing on the creative, museum and educational sectors

– techniques for the automatic recognition of single and multiple touch events and gestures. As well as the gestures commonly used in GUI applications such as clicking and dragging, the work with the Rock demonstrated the possibility of detecting more unusual gestures such as grasping and stroking, which may be of interest for those designing applications which do not fit a standard desktop paradigm.With the addition of a more sophisticated hand tracker, it may be possible to further improve the tracking, by tracking touch actions using hand shape as well as fingertip cues, although it is not currently clear whether these may require per-individual training.

Acknowledgements This work is funded by the EPSRC.

References

1. Synaptics Capacitive sensing technical description. http://www.synaptics.com/technology/tcps.cfm
2. Bérard, F.: The magic table: Computer-vision based augmentation of a whiteboard for creative meetings. In: IEEE workshop on Projector-Camera Systems. (2003)
3. Morrison, G.D.: A cmos camera-based man-machine input device for large-format interactive displays. In: SIGGRAPH. (2007)
4. Malik, S., Laszlo, J.: Visual touchpad: A two-handed gestural input device. In: ACM Int. Conference on Multimodel Interfaces. (2004)
5. Wilson, A.: Touchlight: An imaging touch screen and display for gesture-based interaction. In: Int. Conf. on Multimodal Interfaces. (2004)
6. Wacom Graphics tablets. http://www.wacom.com
7. MagicTouch Touch screens. http://www.magictouch.com
8. Synaptics Touch pads. http://www.synaptics.com
9. Dietz, P., Leigh, D.: Diamondtouch: A multi-user touch technology. In: UIST. (2003)
10. Schmidt, A., Strohbach, M., van Laerhoven, K., Friday, A., Gellersen, H.W.: Context acquisition based on load sensing. In: UbiComp. (2002)
11. Strickon, J., Paradiso, J.: Tracking hands above large interactive surfaces with a low-cost scanning laser rangefinder. In: CHI. (1998)
12. Han, J.Y.: Low-cost multi-touch sensing through frustrated total internal reflection. In: UIST. (2005)
13. Tosas, M., Li, B.: Virtual touch screen for mixed reality. In: Proc. of ECCV Workshop on HCI,. (2004)
14. Benko, H., Wilson, A.D., Baudisch, P.: Precise selection techniques for multi-touch screens. In: CHI. (2006)
15. Wigdor, D., Forlines, C., Baudisch, P., Barnwell, J., Shen, C.: Lucidtouch: A see-through mobile device. In: UIST. (2007)
16. McKenna, S., Gong, S., Raja, Y.: Face recognition in dynamic scenes. In: British Machine Vision Conference. (1997)
17. Yang, J., Fu, Z., Tan, T., Hu, W.: Skin color detection using multiple cues. In: Int. Conference on Pattern Recognition. (2004)
18. Zarit, B.D., Super, B.J., Quek, F.K.H.: Comparison of five colour models in skin pixel classification. In: ICCC International Workshop on recognition, analysis and tracking of faces and gestures in Real-Time systems. (1999)
19. Störring, M., Andersen, H.J., Granum, E.: Skin colour detection under changing lighting conditions. In: 7th Symposium on Intelligent Robotics Systems. (1999)