There's an app for that shirt! Evaluation of augmented reality tracking methods on deformable surfaces for fashion design

Silvia Ruzanka[a], Ben Chang[b], Katherine Behar[c]
[a,b] Rensselaer Polytechnic Institute, 110 8th Street, Troy, NY, USA;
[c]Baruch College, City University of New York, 55 Lexington Avenue, New York, NY, USA

ABSTRACT

In this paper we present *appARel*, a creative research project at the intersection of augmented reality, fashion, and performance art. *appARel* is a mobile augmented reality application that transforms otherwise ordinary garments with 3D animations and modifications. With *appARel*, entire fashion collections can be uploaded in a smartphone application, and "new looks" can be downloaded in a software update. The project will culminate in a performance art fashion show, scheduled for March 2013. *appARel* includes textile designs incorporating fiducial markers, garment designs that incorporate multiple markers with the human body, and iOS and Android apps that apply different augments, or "looks", to a garment. We discuss our philosophy for combining computer-generated and physical objects; and share the challenges we encountered in applying fiduciary markers to the 3D curvatures of the human body.

Keywords: Augmented reality, art, fashion

1. INTRODUCTION

Whether aesthetic or functional, clothing has always served to augment the human body. All clothing extends the biological body's capacities for contact with the natural environment by providing enhanced function such as insulation, waterproofing, sunshade, and ventilation. Garments also disguise, reveal, transform, reshape, and adorn the body, toward aesthetic ends ranging from the everyday to the outlandish.

Garments themselves are generally static objects, but what would happen if they became dynamic – able to change and evolve? Already, textile designers have developed numerous "smart" fabrics ranging from fads like hypercolor to high performance fabrics like sweat-wicking fibers. Through innovations in physical materials, smart fabrics can respond on a functional level to changes in external conditions or in the wearer's bodily state.

appARel uses augmented reality to explore the aesthetic potential of dynamic garments. Augmented reality can enable designers to create increasingly flexible and fluid forms of fashion. By focusing on the aesthetic aspect of dynamic garments, *appARel* allows designers' creativity and artistic vision to surpass the limitations of physical materials.

The history of garment design, both in haute couture and in the genre of contemporary new media art called "wearables," includes examples of garments with mechanical systems, embedded sensors, processors, and displays. Wearables (short for "wearable technology") are clothing or accessories that incorporate and utilize computing or electronic elements. Examples range from solar-powered, motorized, mechanical, or pneumatic clothing like Assa Ashuach's My Trousers (2003) or Ying Gao's Walking City (2006) to garments that respond to environmental conditions like Valérie Lamontagne's Peau d'Âne (2005-2008), to networked communicative accessories that connect multiple people like Katherine Moriwaki and Jonah Brucker-Cohen's Umbrella.net (2004).

While wearables are usually associated with experimental new media art, similar ideas have also been pursued in haute couture. Hussein Chalayan's Fall 2000 and Spring 2007 collections are particularly notable in this regard. Chalayan's Fall 2000 collection consisted of minimal, utopic garments that the models transformed into a set of modernist furniture with a table and chairs, which in turn compacted into suitcases. For his renowned Spring 2007 collection, Chalayan created six robotic "transformer" dresses, each of which slowly morphed through different eras of fashion history, in replete and lavish detail, while his models stood stock-still.

Chalayan is regarded as a provocateur in haute couture; nevertheless, all high fashion is an arena for extravagant expressions of creativity and imagination. Taking an everyday item—clothing—as its basis, haute couture provides a unique intersection between the quotidian and the fantastic. Couture privileges intensive experimentation to transform the essential into the extreme. Fashion caters to this dual role: its products function at once as practical protective coverings and as subjective expressions of self in social relationships and in larger cultural contexts. In this sense, fashion is directly analogous to augmented reality: both exist simultaneously as marker and illusion.

2. PROJECT DESCRIPTION AND CHALLENGES

2.1 DESCRIPTION

Building on this history, *appARel* takes a unique approach, using augmented reality. In this fashion show / performance art project, performers wear garments with prints that serve as visual markers for an augmented reality system. Audience members view the performers through smartphones and tablets running an AR app that augments and transforms the garments.

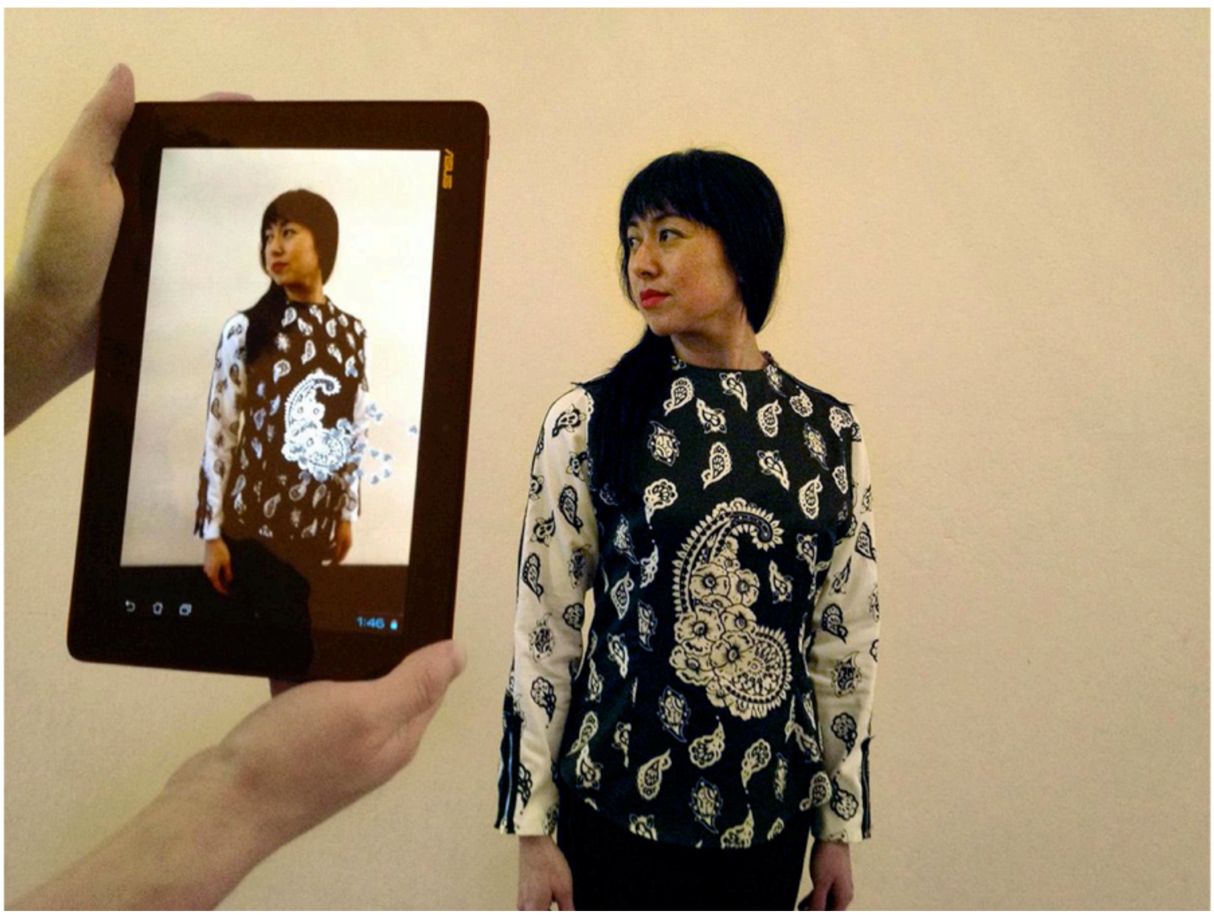

Figure 1: appARel

By combining software and smart devices together with couture, *appARel* makes playful acknowledgement of two other attributes of fashion's cycles. In allowing the visual style of an outfit to be updated with a software refresh, *appARel* accelerates the fashion industry's imperative for perpetual reinvention. At the same time, the physical garment becomes an unchanging canvas for an unlimited number of visual iterations. Rather than issuing a new collection, designers would release an update each season, and instead of investing in new clothes, fashionistas would wear the same garments and simply refresh them with new augments.

2.2 TECHNICAL CHALLENGES

While there are a number of well-established methods for optical AR, this project introduces new challenges for marker tracking. The first is viewing distance. In most AR applications, particularly for mobile devices, the camera is assumed to be close to the marker. For example, placing a marker on a business card, a game piece, a gameboard on a tabletop, or a page in a book assumes that the user will be able to hold the camera as close as needed to provide a clean image for the tracker. For our application, however, users need to be able to stand further back, ideally at least far enough to view a person wearing the augmented garment head to toe. Distance is dictated by several factors. The audience members should be able to see the performer's entire body, to be able to see choreographed movements. Social conventions also dictate a comfort zone of space between peope, so, for example, it would be generally uncomfortable to ask audience members to hold their phones one foot away from the performer's body to view the augments. Performance Art allows this comfort zone to be redefined, but typically in work that involves an explicit breaking of taboos.

The constraint of viewing distance is really based on the size of the marker in the camera view – how many pixels the tracker needs to be able to correctly identify the marker and estimate its pose with an acceptable degree of accuracy. A large marker viewed from a large distance will have same performance as a small marker at a small distance, if both occupy the same pixel area in the camera image. However, we are also constrained by the size of the garment, or more specifically by the size of quasi-planar surfaces on the human body.

Most image-based AR methods assume that the target is on a flat plane. Calculating the homography from the detected target to its original plane provides position and orientation in camera space, which can be used to render the virtual overlay. However, clothing is soft and deformable. Images on a garment might curve across the surface of the body, around the side of the torso, or around a leg or an arm. Folds and wrinkles may also introduce noise, deforming the image target or obscuring parts of it. Likewise, construction processes for shaping a volumetric garment out of a planar textile introduce seams which may crop, obscure, or mismatch portions of a marker.

Our application ideally would allow multiple markers to be tracked simultaneously to enable multiple augments on different parts of the body as well as multiple performers. In our project the marker images are functional elements for the AR system, but must also be incorporated into the garment in a manner that is successful from the perspective of fashion design. Ideally, the functional constraints of the marker design should not overly determine the aesthetic possibilities of the garment and textile design. Finally, since our application targets mobile devices limitations on processing power are a significant factor. Methods which work well on PC hardware may not achieve satisfactory framerate on current-generation mobile platforms.

3. AR METHODS

We examined three approaches: frame-based marker tracking, such as the method used in ARToolkit[1]; color-blob tracking; and feature-based tracking using the proprietary Vuforia (QCAR) library from Qualcomm.

3.1. Frame markers

Frame-based marker tracking assumes that the target image is contained within a square frame. Possible marker candidate areas are warped with an inverse perspective transformation, and their contents identified using one of several methods. After a positive identification, the pose of the marker in camera space can be estimated and used to apply a 3D overlay.

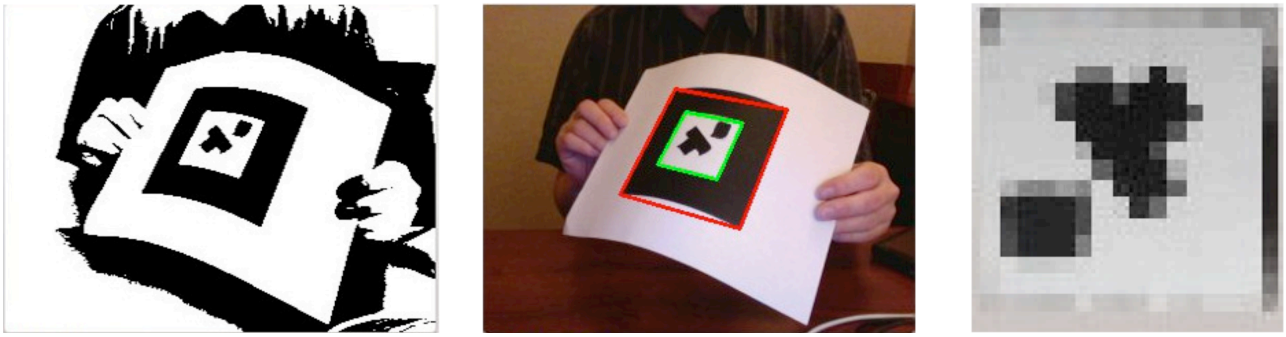

Figure 2: framed marker detection

1. Apply threshold filter
2. Find and simplify contours
3. Keep convex contours with four sides within minimum / maximum dimensions
4. Find inner contours with four sides, indicating that the area is likely to be a frame
5. Use the perspective transform to warp the image of the detected marker area
6. Use template matching to compare against known target images
7. Refine the marker corner positions at the subpixel level
8. Estimate the marker's pose in camera space
9. Draw augmented objects over the camera image

There are a number of variations on this process. Using variable thresholding can reduce problems with changing light conditions, with added computational cost. Alternately, an edge detection filter can be used instead of thresholding. The matching stage can use feature-point matching instead of template matching, which can allow more complex and visually meaningful images within the tracker frame and faster searching with large sets of markers.

ARTag[2] uses encoded images within the rectangular frame. The image is a grid of black and white squares, representing bits encoding a numerical identifier for the marker. Encoding schemes such as Hamming Code can provide error-correction for more accurate identification, in systems such as Aruco[3]. Placing the encoding bits around the interior edge of the frame leaves the majority of the interior space unused by the tracker, so it can be filled instead with visually meaningful imagery.

3.2. Color tracking

One of the most basic methods for object tracking is blob tracking, which detects regions of interest based on color, brightness, or other similar criteria. If the garment is made of a brightly colored fabric, blob tracking might be used to first isolate the garment before further processing. If the garment has a pattern in a second color on a strong background color, the pattern elements can be detected as a second round of blob tracking. This will also reduce the number of false positives that would be generated when just detecting a single color. For example, detecting a blue shirt would also detect other blue objects; but if the shirt is blue with yellow spots, it is less likely that there will be other blue objects with yellow spots.

Figure 3: Polka dot detector. Left to right: original image, primary color mask, dot mask, augmented image

In the context of augmented reality fashion, this process gives us a Polka Dot Detector, to detect objects of color C1 with dots of color C2:

1. Mask image pixels matching C1 within a specified tolerance
2. Find contours in the mask image
3. Discard contours below a given size threshold, fill the remaining contours with white
4. Use the mask image to isolate the garment from the original image. Apply it to a background color which will not be matched by searching for color C2 in the next stage.
5. Mask image pixels matching C2 to detect polka dots
6. Find contours on the new mask image to produce contours around the dots
7. Apply shape heuristics (width/height ratio, convexity test, actual area vs. ellipse area test) to eliminate non-dots
8. Apply augments at dot center points.

Without shape heuristics this method detects any colored shapes on another colored background; using additional criteria it could be adapted to detect squares, triangles, stars, hearts, or other common textile prints. It can detect any number of polka dots, allowing an augment to appear at every dot. However, it cannot distinguish between dots, so will only work well when the same augment needs to be repeated multiple times. It also does not supply enough information to reliably estimate the 3D pose, but 2D scale and position alone may be sufficient either for 2D augments such as images or video, effects produced by filtering or warping the original image, or augments which don't rely on accurate 3D rotation and position such as spheres or particle effects.

3.3. Natural feature-point tracking

Natural feature tracking, sometimes called "markerless" tracking, relies only on keypoints within the image, rather than the morphological constraints of frame-based tracking. This provides far greater flexibility in the kinds of images that can be used for AR markers. Since these methods rely on large numbers of feature points, they in fact tend to work best with complex, irregularly textured images. The computational cost of feature descriptors such as the popular SIFT[4] has presented challenges in their adaptation to realtime applications on mobile phones. One avenue has been to optimize and improve performance of SIFT one mobile hardware.[5] Other efforts to improve peformance with new keypoint descriptors include SURF[6], ORB[7], BRIEF[8], BRISK[9], and FREAK[10].

The continual improvement of both mobile phone hardware and keypoint descriptor algorithms has made AR on phones a widespread possibility, and has led to a number of commercial and proprietary AR libraries for mobile operating systems. Vuforia [11] , formerly known as QCAR, is an AR system produced by Qualcomm which uses natural feature tracking, runs on iOS and Android, and includes a plugin for the popular Unity3D game engine. Vuforia uses a web-based service to analyze target images, detect keypoints and generate descriptor data, which the developer then downloads and bundles into their application. The Unity3D plugin provides a simple interface to manage detection and tracking and lets the developer leverage all the graphical capabilities of the engine (shaders, realtime shadows, skeletal deformers, deferred rendering, particle systems, rigid body dynamics, post-processing effects, etc). As artists, these features made Vuforia particularly attractive for our application, and natural feature tracking allowed us to go beyond the limitations of the black frame.

4. PAISLEY AND PERFORMANCE

4.1 Aesthetic considerations

Augmented reality shows simultaneous views of the "real world" accessible to human vision, and of computer-generated imagery facilitated by computer vision. Recently, the real-world artifacts that enable computer vision systems have developed into a fashion trend with a cachet of its own, known as "The New Aesthetic," a term coined by blogger James Bridle [12] , discussed at the 2012 SXSW conference [13] , and further canonized by author Bruce Sterling [14] . The New Aesthetic prizes intersections of human and machine visuality, by placing (humanist) aesthetic value on the style of imagery that is designed primarily for machines to "see." AR markers intended for computer vision, which humans see only inadvertently, are a prime example of the kind of intersection the New Aesthetic admires. However, from an aesthetic perspective, *appARel*'s design does not seek to prioritize the visual style associated with computer vision. In contrast to the New Aesthetic, we explicitly sought to design for human vision, by developing an "Old Aesthetic" of the handmade and organic.

Our textile designs are based on paisleys, a traditional textile design from the Kasmir region of Asia. The featured shaped in paisley textiles, the boteh, is an ancient motif which is cited as dating to Scythian and Achaemenid art. The earliest surviving examples of the boteh motif are found in silk textiles from the 6th–8th centuries CE from the Akhmim region in Egypt. The ancient boteh shape is found in intricately patterned woven textiles produced in Kashmir as brocade shawls. Kashmiri boteh designs were first imported to Europe in the early 17th century by the British East India Company and became wildly popular. To meet demand, European manufacturers began fabricating imitation Kashmiri shawls and weavers in Paisley, a town in Scotland, were among the top innovators in European weaving and the first to surpass two-color designs in Europe. We selected paisleys because unlike other traditional patterned textiles, such as houndstooth, checks, stripes, polka dots, herringbone, etc., paisleys have a high level of internal intricacy and extensive stylistic variation, which makes them well-suited for natural feature-point tracking.

To further our "old aesthetic," *appARel* privileges organic, irregular forms. When augmented, our garments deform and distort the bodies of their wearers in ways that defy the regularity, symmetry, and proportionality that is generally associated with the human form. While fashion and the fashion industry are known and sometimes criticized for striving to create idealized figures, precedents also exist for non-ideal fashion, particularly from Japanese houses. Rei Kawakubo's work for Comme des Garçon is most seminal in this regard. Kawakubo's Spring 1997 collection "Lumps and Bumps" provides a strong influence for the deformations our augments create. Along with lumps and bumps, *appARel* produces cavities, vortices, and clouds inspired by irregular, organic forms. *appARel* deforms both bodily surfaces, as Kawakubo does physically in her garments, and the textiles themselves, an effect we achieve by combining physical garments with software.

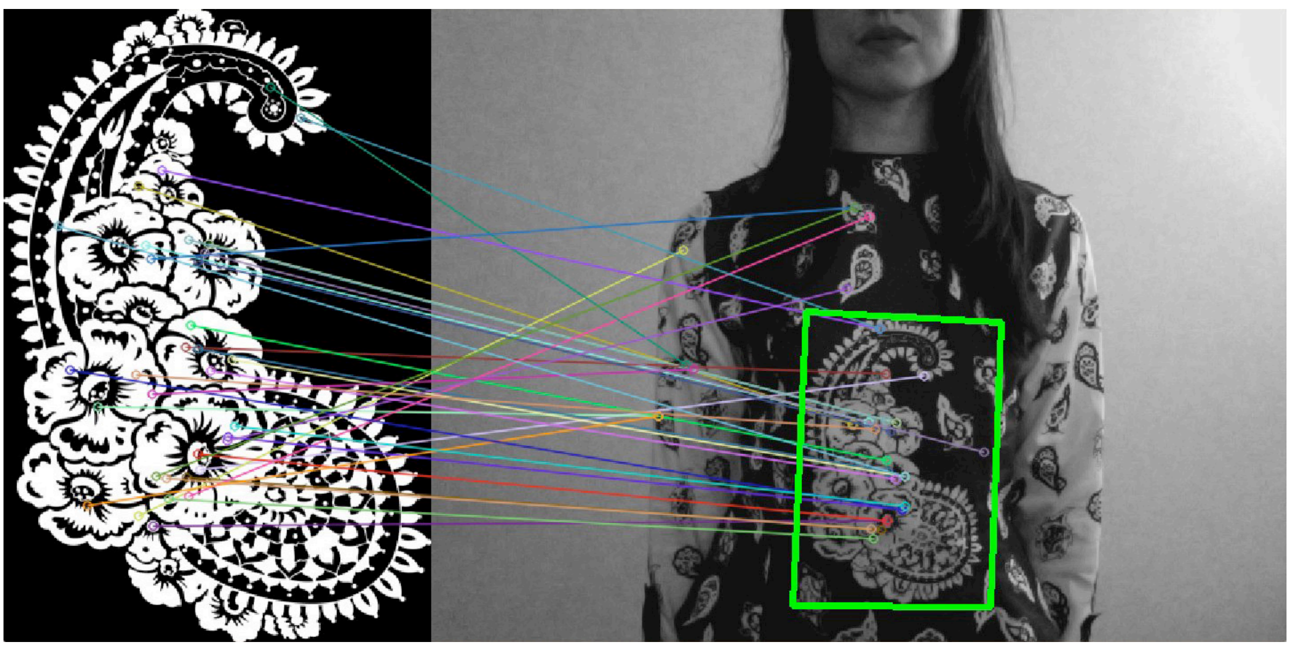

Figure 4: Paisley tracking using OpenCV FREAK

4.2 Textile and garment design

In designing textiles suitable for AR fashion use, we initially sought to use the tiled repeat of the textile to create a mesh of the body, which we could deform through augments. However, to meet the viewing distance and scale challenges described above, our first repeating pattern designs proved either too small to register when viewed through the camera, or clownishly large when applied to the human figure. Additionally, the limited number of markers which can be recognized at a time made using a true repeat impractical for markers. To compensate, we designed textiles that incorporate a single larger marker surrounded by a small-scale repeat where only the larger marker, not the repeat, is tracked.

Next, to achieve the desired effect of isolating and deforming body parts, while compensating for the cropping challenges described above, we strategically combined textile and garment design. After defining the minimum ideal size for a large marker, we identified strategic places on the adult body that provide a surface large enough to accommodate the marker, and are unlikely to bend or deform so severely as to obscure the marker. Surfaces such as the back of the skull, the upper back, the dentrum, etc. proved most practical. We designed our garments to avoid seams in those locations.

Finally, because the textiles are not strict repeats and are tightly constrained with regard to seam placement, it was necessary to customize each textile for a particular garment. At the stage of fashion design known as "pattern making" when a unique three-dimensional garment is reduced to a two-dimensional pattern for use in mass production, we used the pattern pieces as guides to carefully place and size the paisleys in the overall design of the textile yardage. This customization also allowed us to make most efficient use of the yardage, limiting costs.

5. OBSERVATIONS

We measured tracking distance for markers using Vuforia to determine the limitations on size for the markers in our textile prints, and the limitations on audience position for the performance. We used a paisley marker rated with 5 stars ("Excellent tracking quality") in the Vuforia system, and tested tracking response on an HTC EVO 3D smartphone and an Asus Transformer tablet. We measured the maximum distance at which the marker was recognized, and distance at which the marker continues to be tracked before being lost.

Table 1: recognition distance for Vuforia on HTC EVO 3D phone

Marker dimensions	Recognition distance
6" x 9"	74"
5" x 7.5"	64"
4" x 6"	48"
3" x 4.5"	37"

Table 2: recognition distance and tracking distance for Vuforia on Asus Transformer tablet

Marker dimensions	Recognition distance	Tracking distance
6" x 9"	60"	153"
5" x 7.5"	48"	126"
4" x 6"	38"	124"
3" x 4.5"	26"	98"

For comparison, we measured distance for frame markers. We implemented basic frame marker recognition in OpenCV for Android. The test code identifies frames with sufficient size that they could be used for marker identification. For this test we did not implement the complete AR pipeline but focused solely on frame recognition as a "best-case" scenario. The code was implemented without platform optimizations, used one level of pyramid downsampling and adaptive thresholding.

Table 3: Recognition distance for OpenCV-based frame marker recognition on HTC EVO 3D

Marker dimensions	**Recognition distance**
6” x 6”	270”
5” x 5”	200”
4” x 4”	139”
3” x 3”	96”

Using the largest Vuforia marker, audience members will need to be within 5 to 6 feet of the performer, although once a marker is identified the system can maintain tracking at a greater distance. We had initially hoped to use multiple small markers in the textile print to augment multiple points on the body. However, viewing distance would become even more limited. Vuforia is in theory able to track up to five markers simultaneously; however, we were only able to maintain tracking on three consistently, and four at most.

6. CONCLUSION AND FUTURE WORK

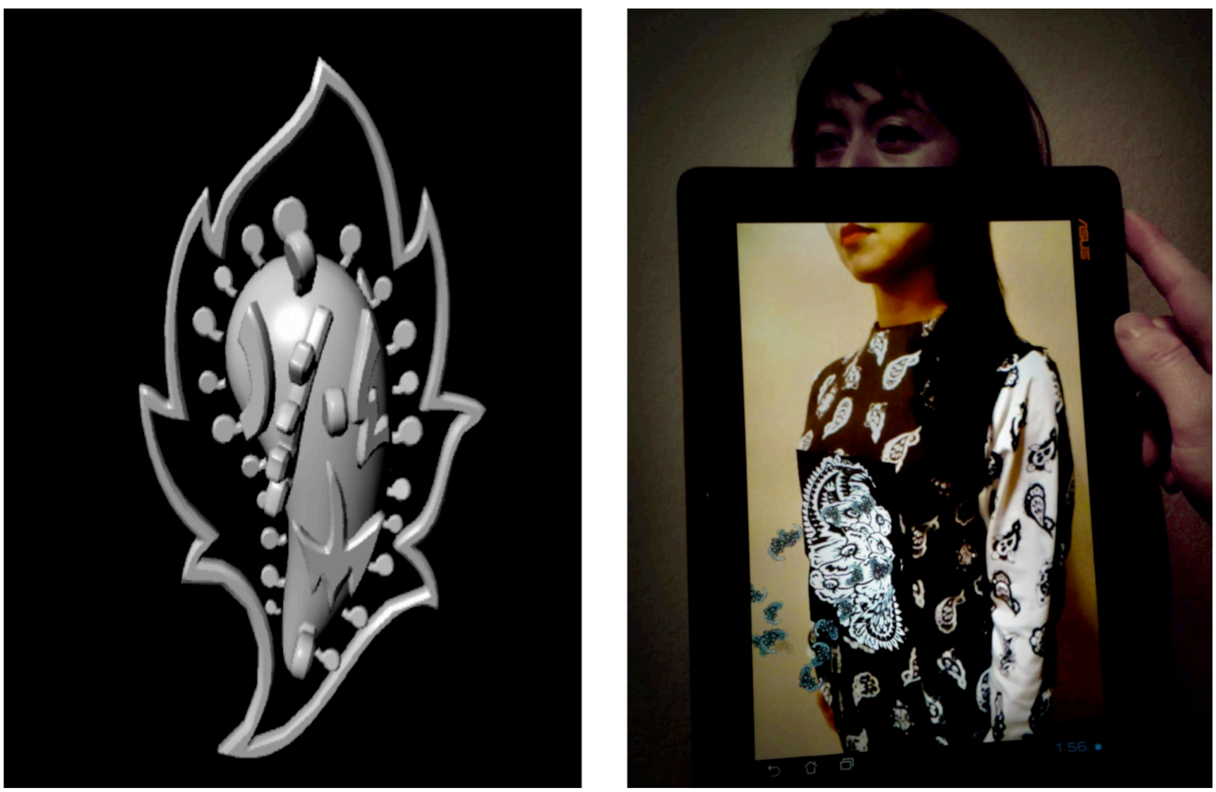

Figure 5: 3D Paisley augment; screenview of *appARel* showing paisley cloud and deformation

appARel is an ongoing art and research project, which will premiere in March 2013 at Monmouth University. The first “collection” of augments include 3D models based on paisley designs; animated abstract structures growing from the body; clouds of paisleys that surround the performer; and augments that deform, extend or collapse parts of the body. The development of the project is a continusing interaction between the possibilities and constraints of media performance, textile design and computer vision. While Vuforia is the current best option for developing our project, the unique characteristics of the problem suggest possibilities in customized and hybrid methods. Frame marker approaches can provide the desired viewing distance and multiple marker tracking, but due to the limitations on textile design we elected to use Vuforia and adapt the performance to its limitations.

We also tested basic feature-based tracking using OpenCV's SURF and FREAK descriptors, to determine whether this would be a viable alternative to Vuforia. Using this approach would allow greater opportunity to customize the overall algorithm and could open up ways to work around some of the limitations built into Vuforia by making other tradeoffs. In initial tests we found that while both SURF and FREAK could reliably track our paisley marker, the distance was not as a good as Vuforia, losing tracking at around two feet. However, there are a number of ways to optimize and fine tune the use of these algorithms, so this warrants further investigation.

ACKNOWLEGEMENTS

Support for this project was provided by a PSC-CUNY Award, jointly funded by The Professional Staff Congress and The City University of New York.

REFERENCES

[1] Kato, H. and Billinghurst, M. Marker Tracking and HMD Calibration for a Video-based Augmented Reality Conferencing System. In *2nd Int'l Workshop on Augmented Reality,* pp 85-94, San Francisco, Ca, Oct 1999.
[2] Fiala, M. "ARTag, An Improved Marker System Based on ARToolkit", NRC/ERB-111, NRC 47166, 2004:36
[3] Muñoz-Salinas, R. "ArUco: a minimal library for Augmented Reality applications based on OpenCv", accessed December 5, 2012. http://www.uco.es/investiga/grupos/ava/node/26
[4] Lowe, D.G. "Distinctive image features from scale-invariant keypoints." International Journal of Computer Vision, 60(2):90-110, 2004.
[5] Wagner, D, et al. "Pose tracking from natural features on mobile phones." Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality. IEEE Computer Society, 2008.
[6] Bay, H., Tuytelaars, T., & Van Gool, L. (2006). Surf: Speeded up robust features. Computer Vision–ECCV 2006, 404-417.
[7] Rublee, Ethan, et al. "ORB: an efficient alternative to SIFT or SURF." Computer Vision (ICCV), 2011 IEEE International Conference on. IEEE, 2011.
[8] Calonder, Michael, et al. "Brief: Binary robust independent elementary features." Computer Vision–ECCV 2010 (2010): 778-792.
[9] Leutenegger, Stefan, Margarita Chli, and Roland Y. Siegwart. "BRISK: Binary robust invariant scalable keypoints." Computer Vision (ICCV), 2011 IEEE International Conference on. IEEE, 2011.
[10] Alahi, Alexandre, Raphael Ortiz, and Pierre Vandergheynst. "FREAK: Fast Retina Keypoint." Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on. IEEE, 2012.
[11] Vuforia (Qualcomm AR): Wagner, Daniel; Barakonyi, Istvan; Siklossy, Istvan; Wright, Jay; Ashok, Roy; Diaz, Serafin; MacIntyre, Blair; Schmalstieg, Dieter; , "Building your vision with Qualcomm's Mobile Augmented Reality (AR) platform: AR on mobile devices," Mixed and Augmented Reality (ISMAR), 2011 10th IEEE International Symposium on , vol., no., pp.1, 26-29 Oct. 2011
[12] Bridle, J. *The New Aesthetic*, http://new-aesthetic.tumblr.com/ (accessed Feb 2, 2013).
[13] SXSW, "The New Aesthetic: Seeing Like Digital Devices." http://schedule.sxsw.com/2012/events/event_IAP11102 (accessed Feb 2, 2013).
[14] Bruce Sterling, "An Essay on the New Aesthetic," *Wired*. April 2, 2012. http://www.wired.com/beyond_the_beyond/2012/04/an-essay-on-the-new-aesthetic/ (accessed Feb 2, 2013).