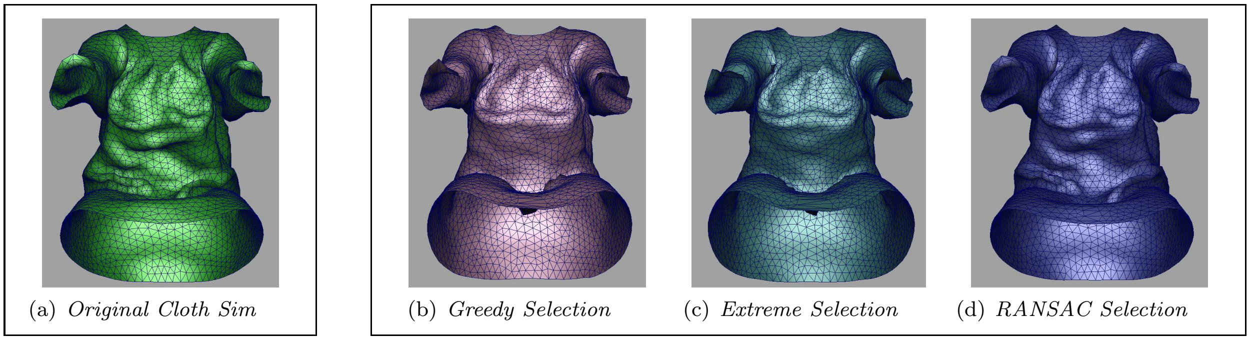
Automated Target Selection for DrivenShape

Gene S. Lee*
Walt Disney Animation Studios



(a) *Original Cloth Sim* (b) *Greedy Selection* (c) *Extreme Selection* (d) *RANSAC Selection*

Figure 1: *Sample frame from an animation that compares different target selection methods for DrivenShape*

1 Introduction

DrivenShape is a data-driven deformation that uses precomputed data (a.k.a. targets) to approximate the effects of a computationally expensive cloth simulation [Kim and Vendrovsky 2008]. Rather than computing a true, accurate solution, DrivenShape produces a quick approximation that satisfies an acceptable margin of error, where error is defined as a difference in appearance or shape.

This paper presents an automated method to select targets for DrivenShape. Our method selects targets that produce the lowest margin of error for DrivenShape in common contexts. Common contexts for DrivenShape include poses, gestures, or configurations that occur frequently or do not change over a period of time. For instance, an animated character usually walks more often than jumps wildly. By focusing on common contexts, this method ensures that deformations of the highest frequencies have low error and the margin of error of DrivenShape is reliably low.

For cleanup artists, who work with DrivenShape, our method allows them to focus on fixing a few frames of high error rather than fixing many frames of average error. This increases their efficiency and speeds up their cleanup times.

2 Target Selection Methods

Two popular methods to select targets for DrivenShape include *extreme posing* and *greedy selection*. With extreme posing, targets are selected to cover extreme configurations. For instance, for legs, these configurations include high kicks and full knee bends. Technical artists commonly apply this method in production, often with great effort. Automated means of extreme posing are rare because the task is subjective and difficult to quantify. With greedy selection, targets are selected to ensure that DrivenShape avoids creating poor approximations. This method usually involves discovering a poor approximation and then finding a target that leads to the best improvement. This method is fast and easy to implement, especially if the number of target candidates is reasonable.

Our method advocates a holistic approach to target selection. It selects targets as a group to reduce error rather than individually selecting them to simply cover poses that are extreme or poorly approximated. The ideal group of targets are those that permit DrivenShape to produce the best approximations in general. Nonetheless, our method does not preclude selecting targets that are either extreme or poorly approximated. If these targets are helpful in reducing error in common contexts, they will ultimately be included.

3 Our Approach

Our method is based upon Random Sample Consensus (RANSAC), a statistical technique for estimating a mathematical model from a data set containing outliers [Fischler and Bolles 1981]. It select targets by repeatedly examining a pool of targets for a valid set of inliers and then finding the model that best fits them. In the context of our work, the inliers define the common contexts and DrivenShape defines the model.

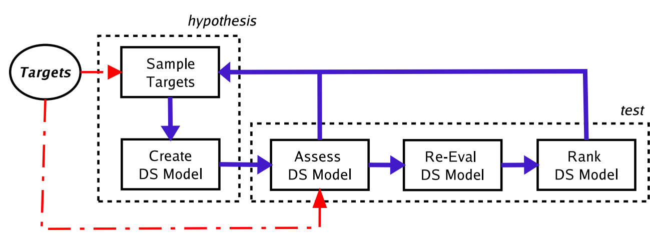


Figure 2: *The workflow for target selection by RANSAC*

The workflow for our method appears in Figure 2. It consists of five steps partitioned into two iterating phases, *hypothesis* and *test*.

3.1 Hypothesis Phase

In the first stage of the hypothesis phase, *Sample Targets*, a small set of targets is randomly selected from a pool of candidates. These samples, which are hypothesized as inliers, can originate from any source, but are best obtained from an animation sequence. Each frame identifies a single target, and the overall sequence provides the data from which to extract a common context. In the second stage, *Create DS Model*, a model is created with DrivenShape to best approximate the samples of the first stage. This is achieved readily by using every sample as a target.

*gene.s.lee@disneyanimation.com
[0]All images ©Disney Enterprises.

3.2 Test Phase

The first stage of the test phase, *Assess DS Model*, assesses the quality of the model from the previous phase. The model qualifies as high quality if it approximates a sufficient number of the remaining targets, which, if extracted from an animation sequence, are the frames that were not selected during the hypothesis phase. The minimum number of targets that must be approximated is a user-defined parameter, which is reasonably set as a percentage of the total number of available targets. The targets approximated by a high-quality model define a set of inliers for the next stage. When the model is a poor approximation, the phase terminates and the process repeats.

The next stage, *Re-Eval DS Model*, re-evaluates the model to approximate the inliers. A proper set of targets is extracted from the inliers using greedy selection. The final stage, *Rank DS Model*, ranks the updated DrivenShape model by assessing its capacity to approximate the inliers. A ranking is ascertained by computing the margin of error for each inlier and then summing the results. The highest ranking model of all iterations provides the best targets for approximating the common context.

4 Results

We applied our approach to a calisthenics pass and several production shots. The calisthenics pass involves a stationary character engaged in an extensive range of motions, while the production shots entail a mobile character undertaking a conventional set of motions. Production environments use a calisthenics pass primarily for review, but also do so for selecting targets by extreme posing.

4.1 Calisthenics Pass

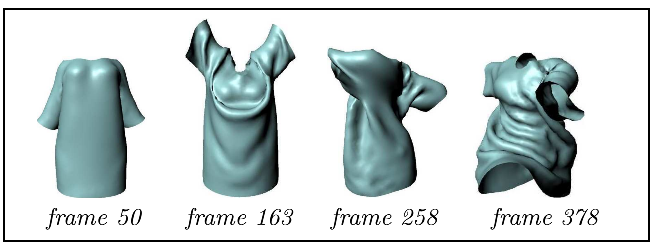


Figure 3: *Sample frames of the Calisthenics Pass*

The calisthenics pass animates a clothed, female character for 566 frames, as shown in Figure 3. An animator sets the animation keys while a proprietary solver simulates the cloth. Each frame of the pass identifies a potential target and the whole pass is a source for extracting a common context.

Target selection on the calisthenics pass was performed three times, with RANSAC, greedy selection (GS), and extreme posing (EP). RANSAC chose targets for the common context with a user-defined parameter of 90% while GS chose targets from frames yielding the highest error. EP, performed manually, chose targets involved in a wide array of postures, starting with the most extreme postures first.

25 Targets	**All Frames**			**Inliner Frames**			**Outlier Frames**		
	Avg	Max	Std	Avg	Max	Std	Avg	Max	Std
Greedy	51.9	103.6	31.3	52.5	103.6	31.3	45.9	101.3	26.5
Extreme	27.3	178.0	36.9	23.0	178.0	34.2	66.0	125.6	37.3
RANSAC	28.0	268.9	43.2	17.3	98.6	22.7	125.7	268.9	59.1

Table 1: *Margin of Error for the Calisthenics Pass*

A sample frame from the three tests appears in Figure 1. Of the three tests, RANSAC produced the best images. Vibrations were less apparent and errors were less noticeable. From a mathematical perspective, as summarized in Table 1, RANSAC introduces greater variance in error. This is to be expected for all frames and especially for outliers. Yet, it is the method that exhibits the lowest margin of error for the inlier frames, which define the common context.

4.2 Production Shot

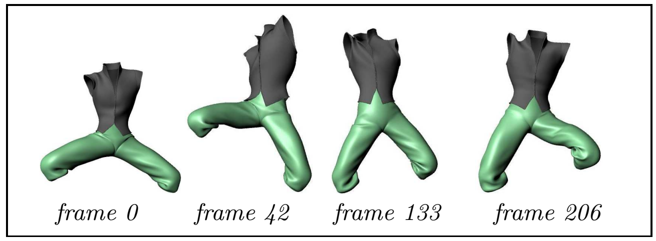


Figure 4: *Sample frames of the Production Shot*

One of the production shots involves a character climbing a wall for 207 frames, as shown in Figure 4. In this shot, the character stops momentarily in his climb to adjust his balance. As before, target selection for the pants, which is approximated by a DrivenShape, was performed three times.

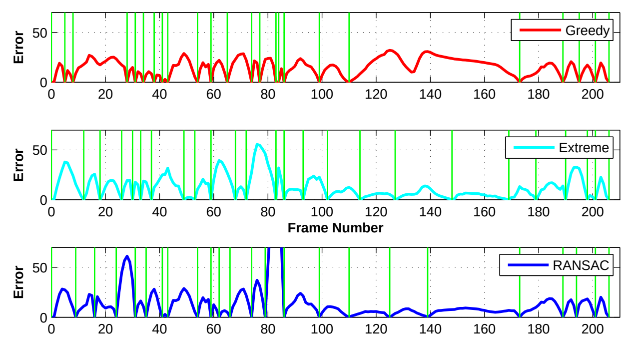


Figure 5: *The error per frame of the climbing shot.*

The visual output of all three methods was similar, yet RANSAC was slightly better. The other methods exhibit greater vibrations and wrinkling at the hips. Figure 5 presents the error per frame for each of the methods. Both RANSAC and EP minimize the error between frames 100 and 180, which is when the character stops momentarily to swivel his hips. With RANSAC, most of the error is present in the outlier frames, 30 and 85. EP distributes the error to all frames before and after the momentary stop while GS distributes the error evenly to every frame. Even with a shot involving a conventional set of motions, RANSAC is capable of extracting an appropriate set of targets.

References

FISCHLER, M. A., AND BOLLES, R. C. 1981. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM 24*, 6, 381–395.

KIM, T.-Y., AND VENDROVSKY, E. 2008. DrivenShape: a data-driven approach for shape deformation. In *SCA '08: Proceedings of the 2008 ACM SIGGRAPH/Eurographics Symposium on Computer Anim ation*, Eurographics Association, Aire-la-Ville, Switzerland, Switzerland, 49–55.