



LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

Comparing Shape and Texture Features for Pattern Recognition in Simulation Data

S. Newsam, C. Kamath

December 13, 2004

IS&T/SPIE's Annual Symposium on Electronic Imaging
San Jose, CA, United States
January 16, 2005 through January 20, 2005

Disclaimer

This document was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor the University of California nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or the University of California, and shall not be used for advertising or product endorsement purposes.

Comparing Shape and Texture Features for Pattern Recognition in Simulation Data

Shawn Newsam and Chandrika Kamath

Lawrence Livermore National Laboratory
7000 East Avenue, Livermore, CA 94550, U.S.A.

ABSTRACT

Shape and texture features have been used for some time for pattern recognition in datasets such as remote sensed imagery, medical imagery, photographs, etc. In this paper, we investigate shape and texture features for pattern recognition in simulation data. In particular, we explore which features are suitable for characterizing regions of interest in images resulting from fluid mixing simulations. Three texture features—gray level co-occurrence matrices, wavelets, and Gabor filters—and two shape features—geometric moments and the angular radial transform—are compared. The features are evaluated using a similarity retrieval framework. Our preliminary results indicate that Gabor filters perform the best among the texture features and the angular radial transform performs the best among the shape features. The feature which performs the best overall is dependent on how the groundtruth dataset is created.

Keywords: Texture features, shape features, simulation data, similarity retrieval

1. INTRODUCTION

Computer simulations are increasingly being seen as the third mode of science, complementing theory and experiments. In order to validate the physics models in these simulations, their results must be compared with an experiment using quantitative measures—this is referred to as “code validation.” A core capability of a system for performing such comparisons is being able to quantitatively characterize the regions of interest in the data. These characterizations enable a host of functionalities such as improved interaction with large datasets, more focused investigations, etc.

Shape and texture features have been used for some time for pattern recognition in datasets such as remote sensed imagery, medical imagery, photographs, etc. In this paper, we investigate shape and texture features for pattern recognition in images resulting from fluid mixing simulations. We consider the problem of shock-driven mixing of two fluids of different densities. When the perturbed interface between the two fluids is accelerated by a shock wave striking the interface perpendicularly, it results in an instability referred to as the Richtmyer-Meshkov instability.^{1,2} This instability occurs in various natural and man-made settings such as supernova explosions, the interiors and wakes of jet engines, combustion chambers, etc.

This work compares how well two different shape features and three different texture features characterize regions of interest in a set of simulation images. For shape features, we consider geometric moments and the angular radial transform. For texture features, we consider grey level co-occurrence matrices, wavelets, and Gabor filters. The features are evaluated using the similarity retrieval framework provided by the similarity based object retrieval (SBOR) system^{3,4} developed by the Sapphire scientific data mining project⁵ in the Center for Applied Scientific Computing at Lawrence Livermore National Laboratory. We consider a dataset of images that consists of two-dimensional slices of the enstrophy variable in a region of a three-dimensional simulation. To mimic the variations between simulations, and within simulations, we include images that are spatial and morphological transformations of the original images. Regions of interest are first identified and used as query regions in the original images. Next, shape and texture features are extracted for the query regions and the regions in the dataset images. These features are used in retrieving regions in the dataset which are similar to the query region. The suitability of the features in characterizing the regions is evaluated by comparing the retrieval results with a groundtruth dataset.

Further author information: S.N.: E-mail: newsam1@llnl.gov

The rest of the paper is organized as follows. Sections 2 and 3 describe the texture and shape features respectively. Section 4 describes how the features are compared including the creation of the groundtruth dataset and the evaluation of the retrieval results. Section 5 presents the results and Section 6 concludes with a summary and ideas for future work.

2. TEXTURE FEATURES

This section describes the three texture features considered in this paper, namely, features based on 1) gray level co-occurrence matrices, 2) wavelets, and 3) Gabor filters.

2.1. Texture Features Based on Gray Level Co-occurrence Matrices

Texture features based on the spatial co-occurrence of pixel values are probably the most widely used in the analysis of remote sensed imagery. First proposed by Haralick⁶ in 1973, they characterize texture using a variety of quantities derived from second order image statistics. Co-occurrence texture features are extracted from an image in two steps. First, the pairwise spatial co-occurrences of pixels separated by a particular angle and/or distance are tabulated using a gray level co-occurrence matrix (GLCM). Second, the GLCM is used to compute a set of scalar quantities that characterize different aspects of the underlying texture. While these quantities can be interpreted using intuitive notions of texture, their main benefit is in providing a quantitative description that can be used for image analysis, such as forming a high-dimensional feature vector to support content based retrieval.

A GLCM is computed for an image as follows. Let $f(x, y)$ be an N_x by N_y image whose pixels take one of L levels, and let \vec{r} be a specified spatial offset. A GLCM entry, $GLCM_{\vec{r}}(i, j)$, is then simply a count of the number of times a pixel with value $j \in 1, \dots, L$ occurs at offset \vec{r} with respect to a pixel with value $i \in 1, \dots, L$:

$$GLCM_{\vec{r}}(i, j) = \#\{(x_1, y_1), (x_2, y_2) \in (N_x, N_y) \times (N_x, N_y) | f(x_1, y_1) = i \wedge f(x_2, y_2) = j \wedge \vec{r} = \overrightarrow{(x_2 - x_1, y_2 - y_1)}\} \quad (1)$$

The offset \vec{r} can be an angle and/or distance. Intuitively, the diagonal and near-diagonal entries of a GLCM will be larger for images composed of patches with the same or similar pixel values, at least with respect to the offset. The off-diagonal entries will be larger for images in which the pixel values vary locally. Particular spatial patterns, or textures, manifest themselves differently in the distribution of entries both within and between GLCMs computed for a variety of offsets.

The GLCM values are typically normalized to sum to one to remove dependence on the image size. Other practical considerations include choosing the number of quantization levels L for images with large dynamic ranges, and whether the quantization should be uniform or incorporate a transformation such as histogram equalization. These choices are dataset dependent.

GLCMs are usually computed for a number of different offsets unless a priori information is available about the underlying texture. A common choice is to compute GLCMs for a distance of one—i.e., adjacency—and four directions, 0, 45, 90, and 135 degrees.

While GLCMs provide a quantitative description of a spatial pattern, they are too unwieldy for practical image analysis. Haralick thus proposed a set of scalar quantities for summarizing the information contained in a GLCM. He originally proposed a total of 14 quantities, or features; however, typically only a subset of these are used. We consider the following five GLCM derived features: angular second moment (ASM), contrast, inverse different moment (IDM), entropy, and correlation. The ASM is computed as

$$ASM = \sum_{i=1}^L \sum_{j=1}^L (GLCM(i, j))^2 \quad (2)$$

It will have a larger value when the co-occurrence frequencies $g(i, j)$ are concentrated in a few places in the GLCM. This often occurs along the diagonal for a constant image or off diagonal for structured images. A near

random or noisy image will have an evenly distributed GLCM with a low ASM. The contrast is computed as

$$\text{CON} = \sum_{n=0}^{L-1} n^2 \left\{ \sum_{|i-j|=n} \text{GLCM}(i, j) \right\}. \quad (3)$$

The contrast is larger for a GLCM with larger off-diagonal values. Thus, contrast is larger for images with quickly varying intensities which agrees with the intuitive notion of contrast. The IDM is computed as

$$\text{IDM} = \sum_{i=1}^L \sum_{j=1}^L \frac{\text{GLCM}(i, j)}{1 + (i - j)^2}. \quad (4)$$

The IDM will be larger for a GLCM with large diagonal values. Thus, the IDM will be large for images with constant or near-constant patches. The entropy is computed as

$$\text{ENT} = - \sum_{i=1}^L \sum_{j=1}^L \text{GLCM}(i, j) \log \text{GLCM}(i, j). \quad (5)$$

The entropy is larger for an image with an evenly distributed GLCM. Thus, a near random or noisy image will have a larger entropy. Finally, the correlation is computed as

$$\text{COR} = \sum_{i=1}^L \sum_{j=1}^L \frac{(ij)\text{GLCM}(i, j) - \mu_{i'}\mu_{j'}}{\sigma_{i'}\sigma_{j'}}, \quad (6)$$

where

$$\mu_{i'} = \sum_{i=1}^L \sum_{j=1}^L i \text{GLCM}(i, j), \quad (7)$$

$$\mu_{j'} = \sum_{i=1}^L \sum_{j=1}^L j \text{GLCM}(i, j), \quad (8)$$

$$\sigma_{i'}^2 = \sum_{i=1}^L \sum_{j=1}^L \text{GLCM}(i, j) (i - \mu_{i'})^2, \text{ and} \quad (9)$$

$$\sigma_{j'}^2 = \sum_{i=1}^L \sum_{j=1}^L \text{GLCM}(i, j) (j - \mu_{j'})^2. \quad (10)$$

A texture feature vector for an image can be formed from the scalar features derived from GLCMs computed for different offsets. For example, a feature vector composed of the above five features computed at four offsets could be formed as

$$h_{\text{GLCM}} = [\text{ASM}_1, \text{CON}_1, \text{IDM}_1, \text{ENT}_1, \text{CORR}_1, \dots, \text{ASM}_4, \text{CON}_4, \text{IDM}_4, \text{ENT}_4, \text{CORR}_4]. \quad (11)$$

2.2. Wavelet Texture Features

Wavelets are not a single image analysis technique but rather a framework for a variety of related techniques. The common appeal of the techniques is that they allow analysis that is localized in both space and frequency. This is accomplished by decomposing an image using bases of small waves, or wavelets, that have compact, sometimes finite, spatial support. Wavelets are used to characterize texture for the same reason as Fourier based techniques—they provide information about the frequency content of an image which is often associated with texture. However, wavelets, unlike Fourier based approaches, are also localized in space.⁷ This makes them more suitable for analyzing texture in non-stationary or non-homogeneous images, such as in remote sensing applications. Wavelets accomplish this not by violating the signal processing analog of the Heisenberg

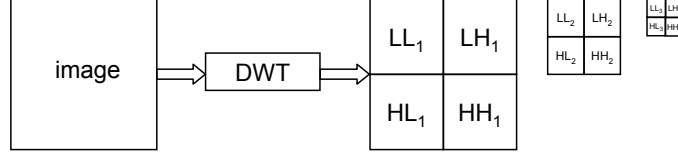


Figure 1. Three level decomposition using the discrete wavelet transform.

uncertainty principle, that analysis cannot be simultaneously localized in space and frequency, but through a multi-resolution analysis (MRA)⁸ that successively trades-off spatial for frequency resolution.⁹

We use the separable discrete wavelet transform (DWT) in an MRA framework to compute the wavelet texture features, as shown in Fig. 1. The DWT applies the Daubechies-4 wavelet to decompose the image into low-low (LL), low-high (LH), high-low (HL), and high-high (HH) components. The MRA framework successively reapplies the DWT to the decimated LL band, resulting in a three level decomposition. The energies and standard deviations of the four components at three resolutions form the texture feature vector:

$$h_{WAVELET} = [energy(LL_1), stdev(LL_1), energy(LH_1), stdev(LH_1), \dots, energy(HH_3), stdev(HH_3)] . \quad (12)$$

Other authors have found that there is very little variation in retrieval performance between texture features derived using different wavelets.¹⁰

2.3. Gabor Texture Features

Texture analysis using filters based on Gabor functions falls into the category of frequency-based approaches. These approaches are based on the premise that texture is an image pattern containing a repetitive structure that can be effectively characterized in a frequency domain, such as the Fourier domain. One of the challenges, however, of such an approach is dealing with the tradeoff between the joint uncertainty in the space and frequency domains. Meaningful frequency based analysis cannot be localized without bound. An attractive mathematical property of Gabor functions is that they minimize the joint uncertainty in space and frequency.¹¹ They achieve the optimal tradeoff between localizing the analysis in the spatial and frequency domains.

Using Gabor filters to analyze texture appeals from a psycho-visual perspective as well. In particular, researchers have found that Gabor functions can be used to model the receptive fields of simple cells in the mammalian visual cortex.¹² Such a finding suggests that Gabor-like filtering takes place early on in the early human visual system.

The texture analysis is accomplished by applying a bank of scale and orientation selective Gabor filters to an image. These filters are constructed as follows.¹³ A two-dimensional Gabor function $g(x, y)$ and its Fourier transform $G(u, v)$ can be written as:

$$g(x, y) = \left(\frac{1}{2\pi\sigma_x\sigma_y} \right) \exp \left[-\frac{1}{2} \left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} \right) + 2\pi i W x \right] \quad (13)$$

and

$$G(u, v) = \exp \left\{ -\frac{1}{2} \left[\frac{(u - W)^2}{\sigma_u^2} + \frac{v^2}{\sigma_v^2} \right] \right\} \quad (14)$$

where $i = \sqrt{-1}$, $\sigma_u = 1/2\pi\sigma_x$ and $\sigma_v = 1/2\pi\sigma_y$ control the tradeoff between spatial and frequency resolution, and W controls the modulation. A class of self-similar functions referred to as Gabor wavelets is now considered. Let $g(x, y)$ be the mother wavelet. A filter dictionary can be obtained by appropriate dilations and translations of $g(x, y)$ through the generating function:

$$\begin{aligned} g_{rs}(x, y) &= a^{-s} g(x', y'), \quad a > 1, \quad s \in 0, \dots, S-1, \quad r \in 1, \dots, R \\ x' &= a^{-s} (x \cos \theta + y \sin \theta) \quad \text{and} \\ y' &= a^{-s} (-x \sin \theta + y \cos \theta) \end{aligned} \quad (15)$$

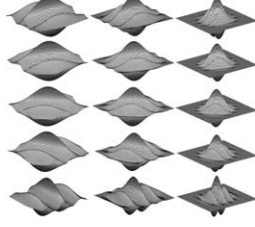


Figure 2. A bank of Gabor filters tuned to combinations of three scales and five orientations.

where $\theta = (r - 1) \pi / R$. The indices r and s indicate the orientation and scale of the filter respectively. R is the total number of orientations and S is the total number of scales in the filter bank. The scale factor a^{-s} in Eq. 15 is meant to ensure that the energy is independent of s . While the size of the filter bank is application dependent, experimentation has shown that a bank of filters tuned to combinations of five scales, at octave intervals, and six orientations, at 30-degree intervals, is sufficient for the analysis of remote sensed imagery. Fig. 2 displays the real components of a bank of Gabor filters tuned to combinations of three scales and five orientations. See Manjunath and Ma¹³ for more details on the filter bank construction, including how the various parameters, such as σ_u , σ_v , a , and W , are derived.

A Gabor texture feature vector is formed from the filter outputs as follows.^{13–15} Applying a bank of Gabor filters with R orientations and S scales to an image results in a total of $R \times S$ filtered images:

$$f'_{11}(x, y), \dots, f'_{RS}(x, y) . \quad (16)$$

The feature vector is formed by computing the first and second moments of the filtered images. That is, an $2RS$ dimension feature vector, h_{GABOR} , is formed as

$$h_{GABOR} = [\mu_{11}, \sigma_{11}, \mu_{12}, \sigma_{12}, \dots, \mu_{1S}, \sigma_{1S}, \dots, \mu_{RS}, \sigma_{RS}] , \quad (17)$$

where μ_{rs} and σ_{rs} are the mean and standard deviation, respectively, of $f'_{rs}(x, y)$.

A homogeneous texture descriptor based on Gabor filters was recently standardized by the MPEG-7 Multimedia Content Description Interface¹⁶ with the minor modification that the mean and standard deviation of the filter outputs are computed in the frequency domain for efficiency.

The Gabor texture features as described above are not rotationally invariant. A simple way to make them approximately rotationally invariant is to circularly shift the feature vectors so that the scale components with the highest mean value are at the first orientation. In this work, we consider these transformed features in addition to the original orientation sensitive ones.

3. SHAPE FEATURES

This section describes the two shape features considered in this paper, namely, features based on 1) the angular radial transform (ART), and 2) geometric moments. Two dimensional shape descriptors typically fall into one of two categories, region-based descriptors and contour-based descriptors. The two features considered in this paper are region-based since they characterize the spatial distribution of pixel intensities. This includes both boundary and interior pixels. A major benefit of region-based shape descriptors is that they can describe complex objects consisting of multiple disconnected regions as well as simple objects with or without holes.¹⁶ In addition, the region-based descriptors considered here do not require the challenging pre-processing step of segmentation.

3.1. ART Shape Features

ART shape features are derived by decomposing an image into a set of orthogonal two-dimensional complex basis functions based on the angular radial transform as follows.¹⁶ Let $f(\rho, \theta)$ be the image in polar coordinates. An ART coefficient of order n and m , $F_{n,m}$, is computed as:

$$F_{n,m} = \int_0^{2\pi} \int_0^1 V_{n,m}^*(\rho, \theta) f(\rho, \theta) \rho d\rho d\theta , \quad (18)$$

where $V_{n,m}(\rho, \theta)$ is the ART basis function that is separable along the angular and radial directions:

$$V_{n,m}(\rho, \theta) = A_m(\theta)R_n(\rho) . \quad (19)$$

A complex exponential function is used for the angular basis function to make the features rotationally invariant:

$$A_m(\theta) = \frac{1}{2\pi} e^{jm\theta} . \quad (20)$$

The radial basis function is a cosine function:

$$R_n(\rho) = \begin{cases} 1 & n = 0 \\ 2\cos(\pi n\rho) & n \neq 0 . \end{cases} \quad (21)$$

The ART shape feature vector is composed of the magnitudes of the complex ART coefficients for twelve angular and three radial basis functions:

$$h_{ART} = [|F_{0,0}|, |F_{0,1}|, \dots, |F_{0,11}|, \dots, |F_{2,11}|] . \quad (22)$$

Finally, each component is normalized by dividing by the first component, $|F_{0,0}|$, which can then be dropped. The ART shape features thus computed are scale and orientation invariant.

3.2. Shape Features Based on Geometric Moments

Shape features based on moments consider an image to be a two-dimensional probability density function. Standard statistical moments can then be used to summarize the spatial distribution of pixel intensities. A number of rotation-, translation-, and scale-invariant values based on these moments then form the shape feature.

The geometric moment shape features considered in this paper are computed as follows.¹⁷ The geometric moment of order $(p + q)$ for image $f(x, y)$ is computed as:

$$m_{p,q} = \sum_{x=0}^{N_x-1} \sum_{y=0}^{N_y-1} x^p y^q f(x, y) . \quad (23)$$

Translation invariance is achieved by using centralized moments:

$$\mu_{p,q} = \sum_{x=0}^{N_x-1} \sum_{y=0}^{N_y-1} (x - x_c)^p (y - y_c)^q f(x, y) \quad (24)$$

where x_c and y_c represent the centroid which can be computed as:

$$x_c = \frac{m_{1,0}}{m_{0,0}} , \quad (25)$$

and,

$$y_c = \frac{m_{0,1}}{m_{0,0}} . \quad (26)$$

The normalized central moments are computed as:

$$\vartheta_{p,q} = \frac{\mu_{p,q}}{(\mu_{0,0})^\gamma} \quad (27)$$

where

$$\gamma = \frac{p+q}{2} + 1 . \quad (28)$$

Finally, seven rotation-, translation-, and scale-invariant values can be computed from the normalized central moments¹⁸:

$$\varphi_1 = \vartheta_{2,0} + \vartheta_{0,2} , \quad (29)$$

$$\varphi_2 = (\vartheta_{2,0} - \vartheta_{0,2})^2 + 4\vartheta_{1,1}^2, \quad (30)$$

$$\varphi_3 = (\vartheta_{3,0} - 3\vartheta_{1,2})^2 + (3\vartheta_{2,1} - \vartheta_{0,3})^2, \quad (31)$$

$$\varphi_4 = (\vartheta_{3,0} + \vartheta_{1,2})^2 + (\vartheta_{2,1} + \vartheta_{0,3})^2, \quad (32)$$

$$\begin{aligned} \varphi_5 = & (\vartheta_{3,0} - 3\vartheta_{1,2})(\vartheta_{3,0} + \vartheta_{1,2}) \left[(\vartheta_{3,0} + \vartheta_{1,2})^2 - 3(\vartheta_{2,1} + \vartheta_{0,3})^2 \right] \\ & + (3\vartheta_{2,1} - \vartheta_{0,3})(\vartheta_{2,1} + \vartheta_{0,3}) \left[3(\vartheta_{3,0} + \vartheta_{1,2})^2 - (\vartheta_{2,1} + \vartheta_{0,3})^2 \right], \end{aligned} \quad (33)$$

$$\varphi_6 = (\vartheta_{2,0} - \vartheta_{0,2}) \left[(\vartheta_{3,0} + \vartheta_{1,2})^2 - (\vartheta_{2,1} + \vartheta_{0,3})^2 \right] + 4\vartheta_{1,1}(\vartheta_{3,0} + \vartheta_{1,2})(\vartheta_{2,1} + \vartheta_{0,3}), \quad (34)$$

and,

$$\begin{aligned} \varphi_7 = & (3\vartheta_{2,1} - \vartheta_{0,3})(\vartheta_{3,0} + \vartheta_{1,2}) \left[(\vartheta_{3,0} + \vartheta_{1,2})^2 - 3(\vartheta_{2,1} + \vartheta_{0,3})^2 \right] \\ & - (\vartheta_{3,0} - 3\vartheta_{1,2})(\vartheta_{2,1} + \vartheta_{0,3}) \left[3(\vartheta_{3,0} + \vartheta_{1,2})^2 - (\vartheta_{2,1} + \vartheta_{0,3})^2 \right]. \end{aligned} \quad (35)$$

The geometric moments shape feature vector is composed of these normalized central moments:

$$h_{GM} = [\varphi_1, \dots, \varphi_7]. \quad (36)$$

4. EXPERIMENTS

4.1. The Original Dataset

For the work in this paper, we consider the data from a high resolution three-dimensional shock tube simulation performed on a $2048 \times 2048 \times 1920$ grid over 27,000 time steps, obtained on 960 nodes of the IBM-SP Sustained Stewardship TeraOp system at Lawrence Livermore National Laboratory.¹⁹ At the beginning of the simulation, two gases are separated by a membrane in a tube. This membrane is then pushed against a wire mesh. The simulation models the resulting mixing of the two gases.

Several variables are output by the simulation for each grid point at each time step. We focus on enstrophy, a byte value with values from 0 to 255. Each compute node in the simulation is responsible for a $256 \times 256 \times 128$ piece of the data. We select nine two-dimensional slices along the wire mesh perpendicular to the direction of impact. These 256×256 sized slices are shown in Figure 3.

Six query tiles of dimension 88×88 are identified in the set of nine images. These tiles are used to evaluate how well the texture and shape features characterize distinctive regions that might be of interest to scientists working with the simulation data. The six tiles are also shown in Figure 3.

4.2. The Extended Dataset

In order to test how robust the features are to spatial transforms, such as rotation and scaling, as well as to minor variations, an extended dataset is generated. A number of spatial and morphological transformations are applied to each of the nine original images. These transformations include:

- Anti-clockwise rotation by 36° , 90° , and 150° .
- Reflection about the vertical axis.
- Decimation and interpolation by a factor of two.
- Morphological erosion and dilation by a 3×3 cross-shaped element.

Examples of these transformations are shown in Figure 4. These transformations are intended to mimic the differences between simulations as well as the differences between simulations and experiments.

The images in the extended dataset are divided into overlapping tiles of dimension 88×88 . The amount of overlap is 80 pixels which results in a target set of 250,470 tiles. The texture and shape features are extracted from each of these target tiles as well as from the six query tiles.

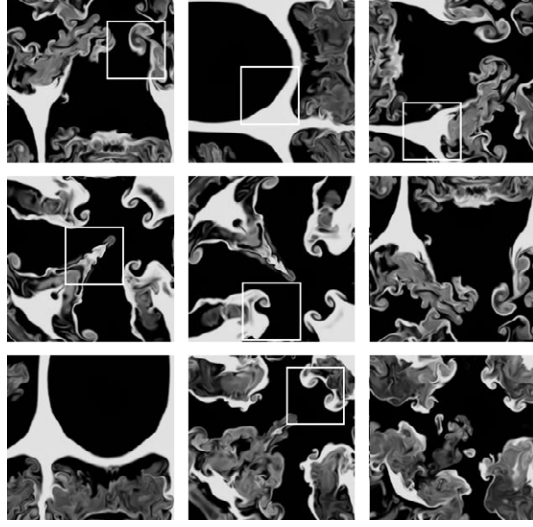


Figure 3. The original dataset consists of nine two-dimensional slices of a high resolution three-dimensional shock tube simulation. The boxes indicate the six query tiles.

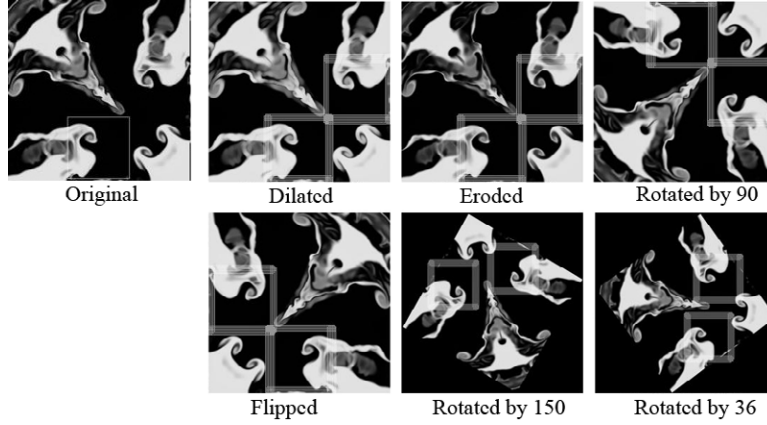


Figure 4. An extended dataset is constructed by applying spatial and morphological transforms to the original dataset. The faint lines indicate the groundtruth for the query region in the original image.

4.3. The Groundtruth Dataset

The groundtruth dataset includes those tiles which are considered the same as, or similar to, a query region. In this work, we construct the groundtruth for a particular query by first identifying other regions in the nine image original dataset that are similar to the query. Then, we determine all the tiles in the extended dataset whose spatial overlap with either the query or one of the similar regions is greater than a specified amount. This is the groundtruth for the query. The faint lines in the transformed images in Figure 4 indicate the groundtruth for the query in the corresponding original image. A threshold of 90% spatial overlap was used in this case. The size of the groundtruth dataset varies from query to query. It ranges from 22 to 81 tiles for a threshold of 90% spatial overlap and from 221 to 421 tiles for a threshold of 75% spatial overlap.

4.4. The Pruned Target Dataset

Finally, tiles that overlap marginally with the query regions are removed from the dataset. For example, all tiles with 50-90% spatial overlap might be removed in the case where the groundtruth consists of tiles with greater than 90% overlap. The final target dataset is then those tiles with 0-40% and 90-100% overlap.

This pruning step creates a more well-defined groundtruth. In particular, it prevents the similarity retrieval framework from favoring those tiles that overlap with the query region in the original image over tiles that

	SHAPE		TEXTURE			
	ART	moments	GLCM	wavlets	Gabor	GaborRI
Feature vector length	35	7	20	24	60	60

Table 1. Feature vector lengths.

overlap with the regions selected as being similar to the query region in the extended dataset. We observed that if this pruning was not performed then tiles with perhaps only 50% overlap with the query region in the original dataset would be considered as more similar to the query than tiles with 95% overlap with one of the similar regions in the extended dataset.

4.5. Average Normalized Mean Retrieval Rate

The texture and shape features are compared using a similarity retrieval framework in which the similarity between a query region and a target tile is computed as the L1 distance between their feature vectors. The basic idea is that the more groundtruth tiles a particular feature retrieves for a query region, the better that feature characterizes the query region.

A k-nearest neighbor (k-NN) search is performed for each of the six query regions in which k is set to twice the number of groundtruth items. A single value, the average normalized mean retrieval rate (ANMRR), is then used to evaluate the retrievals. The ANMRR was originally proposed by the MPEG-7 committee for evaluating visual features.¹⁶ It is computed as follows. Let the groundtruth for query q be $G(q)$. We then perform a k-NN search with k equal to twice the size of $G(q)$. For each feature vector $x \in G(q)$, we define a ranking score $r(x)$ to be the rank of x if x is returned in the k-NN search, or $1.25 \cdot k$ if x is not retrieved at all. We define the normalized modified retrieval rank (NMRR) for query q by averaging the ranking scores for all the feature vectors in $G(q)$, and normalizing the result so that it is within the range of zero and one. Mathematically, NMRR(q) is defined by the following equation:

$$\text{NMRR}(q) = \frac{\sum_{x \in G(q)} r(x) / |G(q)| - 0.5 * (1 + |G(q)|)}{1.25 * k - 0.5 * (1 + |G(q)|)} \quad (37)$$

If the feature vectors in the groundtruth are among the top-ranked results from the k-NN search, NMRR(q) is close to zero. On the other hand, if none of the groundtruth items are retrieved, NMRR(q) becomes one. ANMRR is simply the average of the NMRR scores over all the queries.

5. RESULTS

The ANMRR is computed for each of the following texture and shape features:

- Angular radial transform shape features (ART).
- Geometric moments shape features (moments).
- Grey level co-occurrence texture features (GLCM).
- Wavelet texture features (wavelets).
- Gabor texture features (Gabor).
- Rotationally invariant Gabor texture features (GaborRI).

Table 1 lists the feature vector lengths. High-dimensional indexing methods are typically used to speed up the management of feature vectors in sizable image collections. Since the performance of these methods degrades with increasing feature vector length, a practical application of these features would need to consider their relative vector lengths in addition to how well they characterize regions of interest.

Two experiments are performed:

1. One in which the groundtruth includes those tiles with greater than 90% spatial overlap with the query or similar regions in either the original or extended datasets. In this case, tiles with 50-90% overlap are pruned from the target set.

2. One in which the groundtruth includes those tiles with greater than 75% spatial overlap with the query or similar regions in either the original or extended datasets. In this case, tiles with 50-75% overlap are pruned from the target set.

The ANMRR values for the results of these two experiments are presented in Tables 2 and 3. We make the following observations based on these results:

- The ART shape features outperform the shape features based on geometric moments.
- The shape features based on geometric moments perform poorly.
- The Gabor texture features outperform the GLCM and wavelet texture features.
- The rotationally invariant Gabor texture features outperform the orientation sensitive ones.
- The ART shape features perform the best when the groundtruth overlap is greater than 90% whereas the rotationally invariant Gabor texture features perform the best when the groundtruth overlap is greater than 75%. This suggests that the performance of the Gabor features *degrades more gracefully* as the similarity between the query region and the groundtruth decreases.
- All of the features except for ART seem to degrade gracefully (or even improve) as the similarity between the query regions and the groundtruth decreases.

While the ANMRR metric allows a quantitative comparison of the retrieval performances, visual inspection of the retrieval results allows a qualitative comparison. Figures 6 and 7 show the top 32 retrievals for the ART shape features and rotationally invariant Gabor texture features, respectively, when the groundtruth dataset includes tiles with greater than 90% overlap. The groundtruth dataset for this query is shown in Figure 5. Note that even though the tiles retrieved using the rotationally invariant Gabor texture features are visually more relevant than those retrieved using the ART shape features, the ANMRR value for the Gabor features is actually higher (worse). This is due to the high threshold of 90% spatial overlap. This observation is in keeping with the fact that the ANMRR value for the Gabor features is lower than for the ART features if the threshold is reduced to 75%. Figures 8 and 9 show the top 32 retrievals for the ART shape features and rotationally invariant Gabor texture features, respectively, when the groundtruth dataset includes tiles with greater than 75% overlap.

6. CONCLUSION

This paper investigated the use of a number of different shape and texture features for characterizing regions of interest in simulation data. The features were compared using a similarity retrieval framework in which the result sets were compared against a groundtruth. Using ANMRR as the metric, we observed that the ART features were the best shape features and that the rotationally invariant Gabor features were the best texture features. We also observed that the relative performance of these two best features depended on how the groundtruth was obtained. Future work on this problem includes investigating the use of other quantitative metrics for the comparisons, performing dimension reduction before retrieval, and considering other datasets.

ACKNOWLEDGMENTS

We would like to thank Nu Ai Tang and Sen-ching S. Cheung for their help in developing the SBOR system and in creating the groundtruth.

This work was performed under the auspices of the U.S. Department of Energy by University of California, Lawrence Livermore National Laboratory under Contract W-7405-Eng-48.

	SHAPE		TEXTURE			
	ART	moments	GLCM	wavlets	Gabor	GaborRI
ANMRR	0.52	0.98	0.72	0.73	0.67	0.57

Table 2. ANMRR for the case when the groundtruth includes those tiles with greater than 90% spatial overlap with the query or similar regions in either the original or extended datasets. Tiles with 50-90% overlap are pruned from the target set.

REFERENCES

1. R. D. Richtmyer, "Taylor instability in shock acceleration of compressible fluids," *Communications in Pure and Applied Mathematics* **13**, pp. 297–319, 1960.
2. E. E. Meshkov, "Instability of the interface of two gases accelerated by a shock wave," *Izv. Acad. Sci. USSR Fluid Dynamics* **4**, pp. 101–104, 1969.
3. S.-C. C. E. Cantu-Paz and C. Kamath, "Retrieval of similar objects in simulation data using machine learning techniques," in *SPIE Electronic Imaging. Image Processing: Algorithms and Systems*, 2004.
4. S.-C. Cheung and C. Kamath, "Initial experiences with retrieving similar objects in simulation data," in *Sixth Workshop on Mining Scientific and Engineering Datasets, in conjunction with the Third SIAM conference on Data Mining*, pp. 11–18, 2003.
5. The Sapphire Project on Large Scale Data Mining and Pattern Recognition.
http://www.llnl.gov/CASC/sapphire/sapphire_home.html.
6. R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Transactions on Systems, Man, and Cybernetics* **3**, pp. 610–621, 1973.
7. A. Graps, "An introduction to wavelets," *IEEE Computational Science and Engineering* **2**(2), 1995.
8. S. Mallat, "A theory for multi-resolution signal decomposition: the wavelet representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **11**(7), pp. 674–693, 1989.
9. R. Polikar, "The Wavelet Tutorial." Available at
<http://users.rowan.edu/~polikar/WAVELETS/WTtutorial.html>.
10. W. Y. Ma and B. S. Manjunath, "A comparison of wavelet features for texture annotation," in *Proceedings of the IEEE International Conference on Image Processing*, **2**, pp. 256–259, 1995.
11. J. Daugman, "Complete discrete 2D Gabor transform by neural networks for image analysis and compression," *IEEE Transactions on Acoustics, Speech, and Signal Processing* **36**, pp. 1169–1179, 1988.
12. S. Marcelja, "Mathematical description of the responses of simple cortical cells," *Journal of the Optical Society of America* **70**(11), pp. 1297–1300, 1980.
13. B. S. Manjunath and W. Y. Ma, "Texture features for browsing and retrieval of image data," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **18**(8), pp. 837–842, 1996.
14. P. Wu, B. S. Manjunath, S. D. Newsam, and H. D. Shin, "A texture descriptor for image retrieval and browsing," in *IEEE Conference on Computer Vision and Pattern Recognition: Workshop on Content-Based Access of Image and Video Libraries*, pp. 3–7, 1999.
15. P. Wu, B. S. Manjunath, S. Newsam, and H. D. Shin, "A texture descriptor for browsing and image retrieval," *Journal of Signal Processing: Image Communication* **16**(1), pp. 33–43, 2000.
16. B. S. Manjunath, P. Salembier, and T. Sikora, eds., *Introduction to MPEG7: Multimedia Content Description Interface*, John Wiley & Sons, 2002.
17. M. Sonka, V. Hlavac, and R. Boyle, *Image Processing, Analysis, and Machine Vision*, PWS Publishing, second ed., 1999.
18. M. Hu, "Visual pattern recognition by moment invariants," *IRE Transactions on Information Theory* **IT-8**, pp. 179–187, 1962.
19. A. Mirin et al., "Very high resolution simulation of compressible turbulence on the IBM-SP system," Tech. Rep. UCRL-JC-134237, Lawrence Livermore National Laboratory, 1999.

	SHAPE		TEXTURE			
	ART	moments	GLCM	wavlets	Gabor	GaborRI
ANMRR	0.77	0.97	0.77	0.71	0.65	0.59

Table 3. ANMRR for the case when the groundtruth includes those tiles with greater than 75% spatial overlap with the query or similar regions in either the original or extended datasets. Tiles with 50-75% overlap are pruned from the target set.

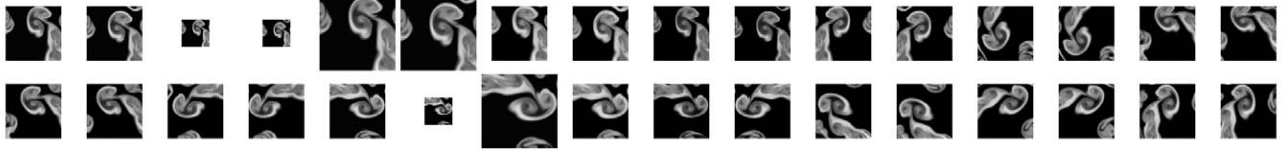


Figure 5. The groundtruth dataset for the first query region for the case when tiles with greater than 90% spatial overlap are included.



Figure 6. The top 32 retrievals for the ART shape features when the groundtruth dataset includes tiles with greater than 90% overlap and tiles with 50-90% overlap are removed from the target dataset.

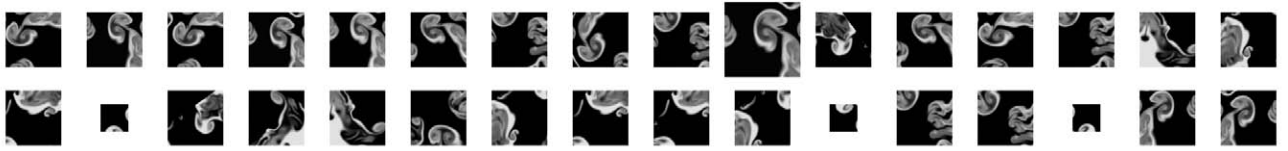


Figure 7. The top 32 retrievals for the rotationally invariant Gabor texture features when the groundtruth dataset includes tiles with greater than 90% overlap and tiles with 50-90% overlap are removed from the target dataset.



Figure 8. The top 32 retrievals for the ART shape features when the groundtruth dataset includes tiles with greater than 75% overlap and tiles with 50-75% overlap are removed from the target dataset.

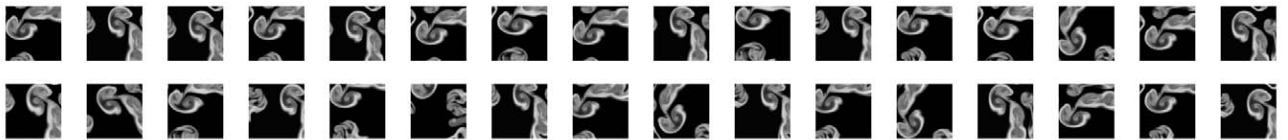


Figure 9. The top 32 retrievals for the rotationally invariant Gabor texture features when the groundtruth dataset includes tiles with greater than 75% overlap and tiles with 50-75% overlap are removed from the target dataset.