Cytometry Part A 64A:101–109 (2005)

Technical Note

Normalization of Multicolor Fluorescence In Situ Hybridization (M-FISH) Images for Improving Color Karyotyping

Yu-Ping Wang[1]* and Kenneth R. Castleman[2]

[1]School of Computing and Engineering, University of Missouri, Kansas City, Missouri
[2]Advanced Digital Imaging Research, LLC, League City, Texas

Received 10 May 2004; Revision Received 29 October 2004; Accepted 15 December 2004

Background: Multiplex or multicolor fluorescence in situ hybridization (M-FISH) is a recently developed cytogenetic technique for cancer diagnosis and research on genetic disorders. By simultaneously viewing the multiply labeled specimens in different color channels, M-FISH facilitates the detection of subtle chromosomal aberrations. The success of this technique largely depends on the accuracy of pixel classification (color karyotyping). Improvements in classifier performance would allow the elucidation of more complex and more subtle chromosomal rearrangements. Normalization of M-FISH images has a significant effect on the accuracy of classification. In particular, misalignment or misregistration across multiple channels seriously affects classification accuracy. Image normalization, including automated registration, must be done before pixel classification.

Methods and Results: We studied several image normalization approaches that affect image classification. In particular, we developed an automated registration technique to correct misalignment across the different fluor images (caused by chromatic aberration and other factors). This new registration algorithm is based on wavelets and spline approximations that have computational advantages and improved accuracy. To evaluate the performance improvement brought about by these data normalization approaches, we used the downstream pixel classification accuracy as a measurement. A Bayesian classifier assumed that each of 24 chromosome classes had a normal probability distribution. The effects that this registration and other normalization steps have on subsequent classification accuracy were evaluated on a comprehensive M-FISH database established by Advanced Digital Imaging Research (http://www.adires.com/05/Project/MFISH_DB/MFISH_DB.shtml).

Conclusions: Pixel misclassification errors result from different factors. These include uneven hybridization, spectral overlap among fluors, and image misregistration. Effective preprocessing of M-FISH images can decrease the effects of those factors and thereby increase pixel classification accuracy. The data normalization steps described in this report, such as image registration and background flattening, can significantly improve subsequent classification accuracy. An improved classifier in turn would allow subtle DNA rearrangements to be identified in genetic diagnosis and cancer research. © 2005 Wiley-Liss, Inc.

Key terms: classification; registration; translocation; wavelets; chromosome image analysis; fluorescence in situ hybridization; automated cytogenetics

The progress of cytogenetics has become significantly faster since the discovery of chromosome banding and molecular fluorescence in situ hybridization (FISH) techniques (1–8). FISH-based methodologies greatly enhance the detection of genetic abnormalities in normal and cancer cells. Multiplex FISH (M-FISH) of human chromosomes is based on the simultaneous hybridization of a pool of 24 chromosome-specific probes (3). Figure 1 is an M-FISH image dataset that is collected from six spectral channels by using a filter wheel. From this multiple channel image, it is able to distinguish each human chromosome in a cell by means of specific color labeling. Figure 2a shows a metaphase chromosome image representing the 24

Contract grant sponsor: National Institutes of Health; Contract grant number: SBIR 2R44HD038151-02.

*Correspondence to: Wang Yu-Ping, Ph.D., School of Computing and Engineering, University of Missouri, 550D RHFH, SCE, 5100 Rockhill Road, Kansas City, MO 64110.

E-mail: wangyup@umkc.edu

Published online 23 February 2005 in Wiley InterScience (www.interscience.wiley.com).

DOI: 10.1002/cyto.a.20116

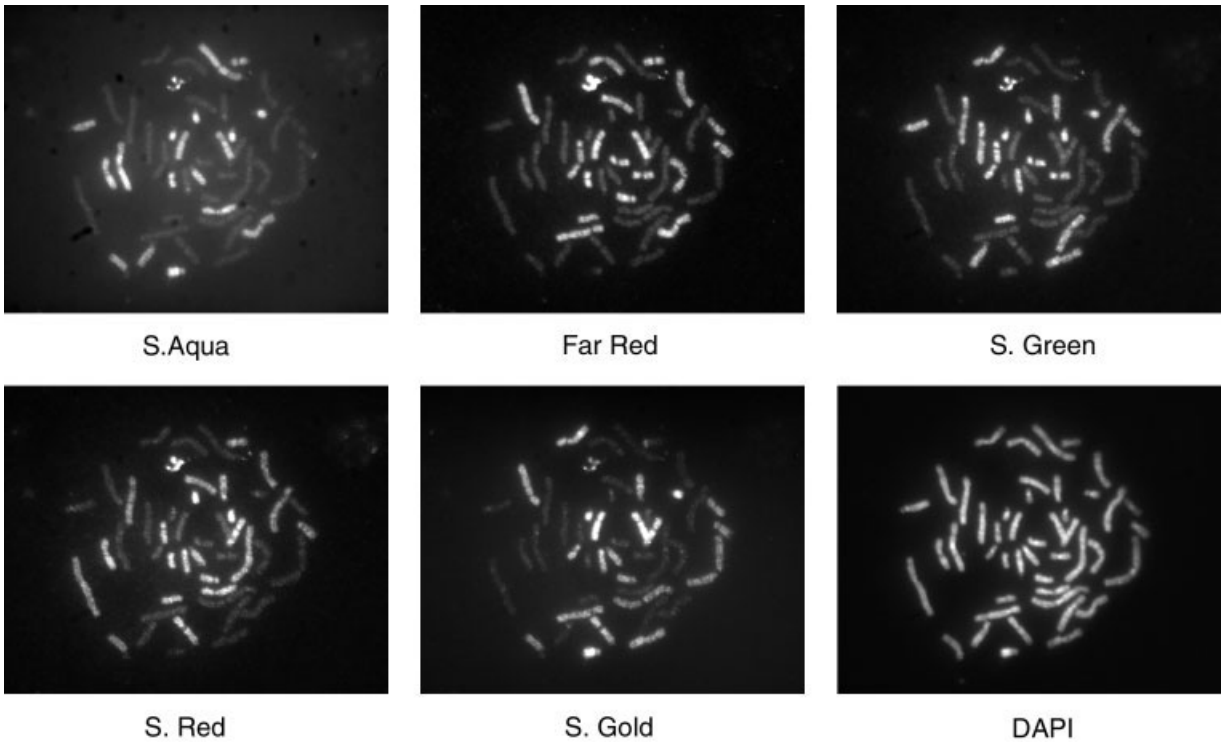

FIG. 1. A six-channel M-FISH image dataset. All chromosomes are stained in the DAPI channel and labeled with a different combination of fluorescence dyes.

classes of chromosomes in a color-mapped display. By using combinations of different colored DNA probes, one can visualize simultaneously the 22 autosomes and the two sex chromosomes with a classifier, as illustrated in Figure 2b. This color karyotyping allows rapid detection of simple and complex chromosomal alterations. The technique can be used for the identification of subtle chromosomal aberrations, such as the translocation of telemetric chromatin, which is difficult to detect with banding alone. It can also be used for identification of small genetic markers that remain elusive after banding.

Digital image processing automates the time-consuming tasks of color karyotyping (1,6,9,10). It provides accurate pixel-by-pixel classification of multicolor FISH images, making M-FISH feasible for many clinical and research applications. Despite its success, the technique has less than perfect classification accuracy (11,12) that is limited by several factors. Figure 2b shows an example of pixel classification with a Bayesian algorithm, which exhibits misclassification of many pixel regions. The size of these misclassified regions is significant compared with those of the small regions involved in the complex chromosomal rearrangements that researchers often hope to elucidate with M-FISH. They also degrade the overall usefulness of the technique. Hence, it is highly desirable that the pixel classification be accurate enough for reliable patient diagnosis. However, even for normal chromosomes the accuracy can not reach 100%. To make this technique practical for identifying chromosome abnormalities in cancer and genetic disease diagnosis, the key step is to increase the classification accuracy. This report describes several data normalization approaches that can significantly affect the accuracy of downstream pixel classification. To evaluate the performance of these approaches, we simply used Bayesian classifier as a test.

There are many factors that can affect a pixel classifier's accuracy. We have found that the discriminatory power (signal-to-noise ratio) and image capture efficiency (excitation efficiency) of different M-FISH analysis systems are useful in measuring a system's ability to obtain accurate and reproducible classification results (12). Misregistration is another important factor that affects the accuracy of pixel classification. In this report, we introduce a sophisticated wavelet-based algorithm for M-FISH image registration. The wavelet technique can improve accuracy and computational speed. To demonstrate the usefulness of registration and other normalization approaches, they are tested by the downstream pixel classification against a public M-FISH database (13) that we previously established.

MATERIALS AND METHODS

We first introduce the database that we established. Then we introduce the several data normalization approaches such as background subtraction and dimension deduction. In particular, a wavelet-based registration algorithm is described in detail. The Bayesian classifier is introduced, which is used to test the performance of data normalization.

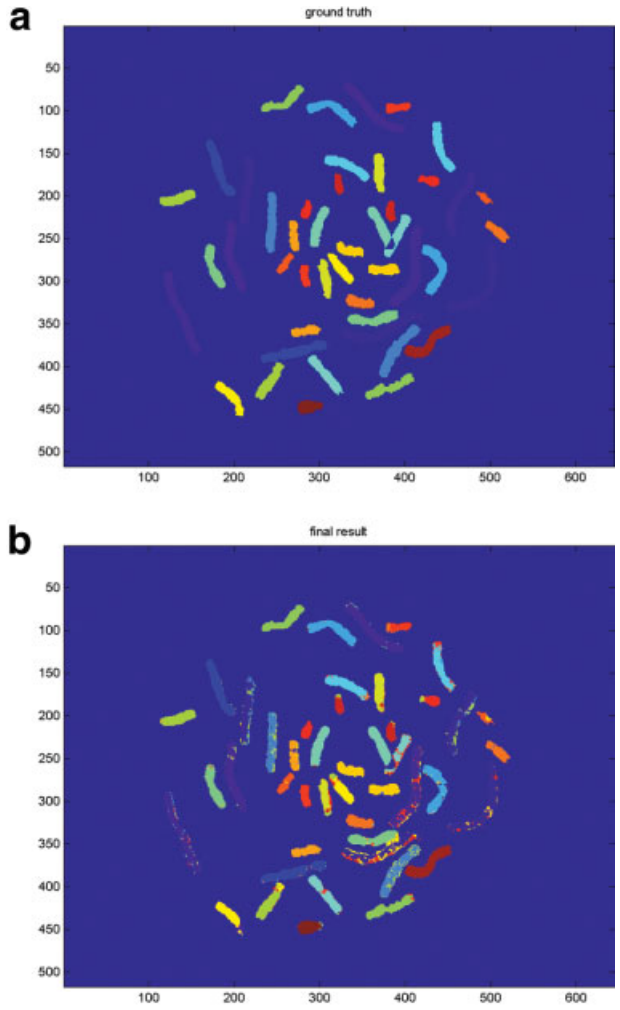

FIG. 2. The real pixel-by-pixel classification of chromosomes from the M-FISH image dataset shown in Figure 1. **a:** Ideal case. **b:** Classification with a Bayesian classifier. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

M-FISH Image Database

A database consisting of 200 M-FISH–labeled human chromosome spread images has been established by Advanced Digital Imaging Research (ADIR) (13–15) to support this research. The database contains six-channel image sets recorded at different wavelengths. The specimens were prepared with probe sets from Applied Spectral Imaging (Migdal HaEmek, Israel), Advanced Digital Imaging Research (ADIR; League City, Texas, USA), Cytocell Technologies (Cambridge, UK), and Vysis (Downers Grove, IL, USA). The database contains 200 spreads from 33 slides from five different laboratories. The specimens include 74 normal male spreads, 8 normal female spreads, 99 abnormal spreads, and 17 more that are of low specimen quality. There are 50 different chromosomal aberrations represented, including numerical abnormalities and structural arrangements. Spread quality ranges from excellent to very difficult. This comprehensive image database is a valuable source for M-FISH studies. In addition, the database includes a classification map, stored as an image file, that was established by experienced cytogeneticists. This image is labeled so that the gray level of each pixel represents its class number (chromosome type). In addition, background pixels are 0, and pixels in a region of overlap are −1. This data file serves as ground truth to test the accuracy of M-FISH image classification algorithms.

In this report, because the goal is to measure the performance of normalization using downstream classification, we use normal chromosome images in which the ground truth is easily established by an expert cytogeneticist. It is reasonable to assume that an improved pixel classifier for normal chromosomes in turn would improve the capability of this technique in resolving subtle and cryptic chromosome abnormalities when used for cancer and genetic diagnosis.

Segmentation of Chromosome Regions

Pixel classification is performed in the region of chromosomes. The first step is to segment the chromosome regions. Because all the chromosomes are stained in the 4′,6-diamidino-2-phenylindole (DAPI) channel and can be visualized under a microscope, the segmentation was done for the DAPI image. First the edges of the chromosomes in the DAPI image are detected using the Laplacian of Gaussian edge detector (16). To fill holes in the edge map, mathematical morphology operations such as dilation and erosion are applied (16) (Fig. 3).

Background Flattening

Intensity variations throughout the image, due to nonuniform illumination and variation among different hybridization experiments, can limit pixel classification accuracy. Background correction or subtraction is necessary to decrease such variations and enhance the fluorescent signals. One approach is to perform background flattening. To do this, we first identify the background points by a segmentation mask. Then we fit a cubic surface to the selected background points (16). The background correction is obtained by subtracting the fitted surface from the image. This normalization procedure turns out to be very important in terms of ultimate classifier performance because it helps to increase the separation distances among the 24 clusters in feature space. This is demonstrated in the Results section.

Dimension Reduction and Feature Selection

Each pixel is defined by its six-dimensional spectra; therefore, each chromosome class is a point in a six-dimensional feature space. A classifier thus has to separate 24 clusters of points in the six-dimensional feature space. The pixel classification accuracy is then limited by the degree of overlap among the 24 clusters of points. As overlapping increases, classification accuracy decreases. Color compensation is a useful preprocessing step. It decreases the effects of spectral overlap among the fluors and in the image sensor. It has been described previously (9,16,17).

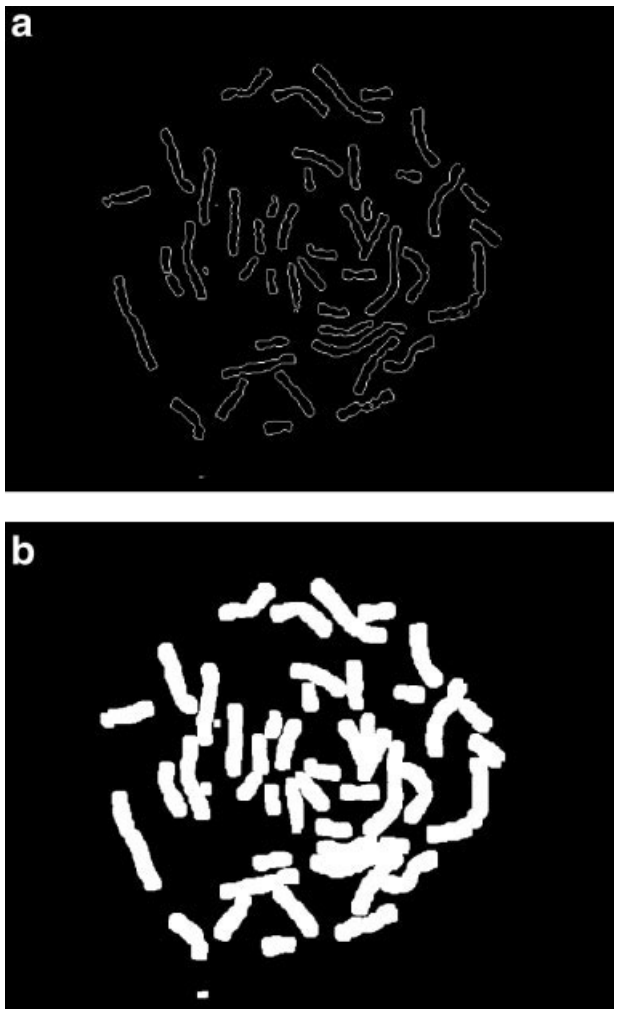

FIG. 3. Segmentation of chromosome regions. **a:** Chromosome boundary detected with Laplacian of Gaussian. **b:** Chromosome region detected followed by mathematical morphology.

Transforming the feature vectors for each class into a new, lower-dimensional feature space might increase the separation distance between clusters. Dimension reduction using principal component analysis was used to map each feature vector into a new, lower-dimensional vector, representing each chromosome in a compact and easily interpretable form. This could eliminate the redundancy between multiple channels that results from spectral overlap.

A vector-based coordinate transform, such as the spherical coordinate representation, is another type of feature space transformation. This may produce better features because only the vector length (not its direction) is affected by the intensity variations that result from nonuniform illumination.

Multispectral Registration

An important factor that limits pixel classification accuracy is the issue of multispectral image registration. An accurate automated registration algorithm is necessary to ensure the accuracy of subsequent pixel classification.

Sources of Misalignment

One optical principle underlying the inherent misregistration problem in M-FISH imaging is that the focal length of the objective lens changes with wavelength. As a consequence, when multiple color filters are used in M-FISH imaging, there is an offset between focal planes of different wavelengths that results in axial chromatic aberration. Moreover, the magnification is inversely proportional to the focal length. Hence, each object in the specimen also changes its off-axis position when imaged with different wavelengths, resulting in lateral chromatic aberration. Many studies conducted on this kind of error have demonstrated that even the best available objectives with high numerical aperture produce axial and lateral chromatic aberration (18). For M-FISH imaging instruments, the effect of chromatic aberration becomes even more severe. Substituting filters in the optical path also causes image shift because it is impossible to maintain perfect mechanical alignment. Aside from these effects, other factors that may cause misregistration problems include vibration of a filter, a lens, a camera, the microscope setup, alignment of the objective, additional optical components plugged into the microscope, and conditions of image acquisition such as changes in ambient temperature, thickness of the coverglass, refractive index of the medium, among others. All of these factors make misregistration unavoidable (Table 1).

Multiresolution Registration

Image deformations. The M-FISH images collected in different color channels must be aligned with each other before further processing, visualization, and karyotyping. We begin with a discussion of two-channel image registration, in which we refer to the reference image and the test image. For an M-FISH image set, the DAPI channel is usually taken as the reference image, and the other five channels are registered to it. The first step of image registration is to find homologous (corresponding) points in both images. Then one must find the geometric transformation, T', between the two images such that they become most similar, in a certain metric, after the transformation.

$$T' = \arg_T \min F(u(\cdot)), v(T(\cdot)) \quad (1)$$

where T' and T represent the geometric transformations; F is the objective function, which measures the similarity between two objects; and $v(T[\cdot])$ stands for the pixel of an image to be registered and is related to the corresponding pixel of the reference image, u, by transformation T. Once a transformation, T', is found, the test image can be warped to the reference image by interpolation. After image warping, the test image should be well aligned with the reference image.

There are generally two classes of image transformations: rigid and nonrigid. For a rigid transformation, trans-

Table 1
List of Sources of Misregistration and the Effects

Name	Causes	Effects
Axial chromatic aberration	Focal length of lens and point spread function changes with wavelength	Axial shift and defocus
Lateral chromatic aberration	Magnification is inversely proportional to the focal length	Lateral off-axis shift and magnification
Mechanical misalignment	Limitations of design, operator actions, optical setup	Shift, rotation, and scaling
Monochromatic aberration	Spherical aberration, curvature of field, astigmatism, distortion, and coma	May lead to nonrigid motion
Other sources	Change of ambient temperature, thickness of cover glass, embedding media, optical components plugged in	May lead to nonrigid motion

lation, rotation, and scaling are allowed. In this report, we have assumed that the geometric distortion is caused by rotation and translation only because the scaling factor is found not to be significant. For an elastic or curved transformation, there can be deformations of the image parts. These deformations are usually described by the local vector displacement (disparity) field. In the case of M-FISH imaging, the transformation is usually assumed to be rigid. Nonrigid transformations could distort the shape of chromosomes and contribute to misdiagnosis.

Similarity criteria. When searching for the transformation, T', in equation (1), there are several objective functions that can be used in the optimization. These including the least square error (LSE), sum of absolute valued differences, and normalized correlation coefficient. The definitions of these metrics are provided elsewhere (19). We have tested several of these metrics and identified the mutual information (MI) metric as most appropriate for M-FISH images (see Results). This metric is defined as

$$MI(T) = \sum_{x} P_{u,v}(x, T[x]) \cdot \log \frac{P_{u,v}(x, T[x])}{P_u(x) \cdot P_v(T[x])} \quad (2)$$

where P_u, P_v, and $P_{u,v}$ are the probability distribution functions of u, v, and their joint probability distribution function. This metric appears to be more suitable for the multispectral images because it makes use of the probability distribution of the gray-level intensities (the histogram) rather than of the image intensities themselves.

Optimization algorithm. The optimization in equation (1) leads to a nonlinear multivariant optimization problem; hence, numerical optimization techniques are required to find the solution. The simplest optimization technique is the full search, which requires division of search space into discrete steps and evaluation of the objective function over this entire set of discrete parameters. Because this technique is time consuming, we choose a more efficient optimization method, depending on the size and structure of the problem. The most popular choices include the Powell method, gradient descent, conjugated gradients, and variations of the Newton method (20,21). We use the Marquardt-Levenberg algorithm in this work.

Image warping or interpolation. After the transformation T' is known, we perform geometric interpolation to unwarp the test image to match the reference image. Because of good approximation properties, spline interpolation is applied (22). This can result in subpixel accuracy.

Multiresolution searching. In applying the abovementioned multidimensional optimization algorithm, convergence speed is a major concern. These algorithms use an iterative strategy and must evaluate derivatives of the objective functions at each step. Using a multiresolution pyramid approach can speed up the process by searching at a lower resolution level. It can also increase robustness through approaching the solution by gradual refinement. Image registration starts at the coarsest level, which has a minimal number of pixels, and then uses that solution as an initial guess to proceed down to the images at finer levels. Optimal transformation parameters are searched only in the neighborhoods of the minima found at the previous level. This is repeated until the finest (original) level is reached. This will decrease searching time. In addition, the multiresolution strategy decreases the sensitivity of the solution to the choice of initial values and increases the accuracy of estimation (23). This is because, at the coarse scale, we can get a close estimation of the transformation parameter rather than a random assignment. These will be demonstrated with examples in the Results.

Evaluation of the normalization approaches using pixel classification. We will use the downstream classification to measure and validate the performance of the data normalization because the ultimate goal of these approaches is to improve the pixel classification accuracy. For this purpose, we simply employed the widely used Bayesian classifier. We have also tested a variety of other classifiers such as the Nearest Neighbor (NN) and K-Nearest Neighbor (kNN) (24) and fuzzy c-means algorithms (25).

Each pixel in the metaphase chromosome image corresponds to a vector with six elements

Table 2
*Comparison of Pixel-by-Pixel Classification of M-FISH Image Sets With and Without Background Correction**

Test image set	Without background flattening	With background flattening
1	75.53	90.87
2	71.92	89.89
3	89.40	93.32
4	90.56	90.42
5	90.25	92.24
Average	83.53	91.35

*Overall correct pixel classification rate is presented as percentage. The experiment was performed on the M-FISH database (13).

$$X = (x_1, x_2, x_3, x_4, x_5, x_6)' \tag{3}$$

where each element of the vector represents the gray-level value in one of the six image channels. This six-dimensional feature vector is fed into a trained pixel classifier, and each pixel in each chromosome is classified into a chromosome type and assigned a corresponding color from a user-defined color table.

Based on the measured six-dimensional features, each pixel in the metaphase image is classified into one of 24 types. Given the nature of data derived from well-separated spectral signatures, the probability distribution of the feature values for each class ω_i is assumed to be multivariate normal density function, with a probability density function as

$$p(X|\omega_i) = \frac{1}{(2\pi)^3 |\sum_i|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}[X-\mu]' \sum_i^{-1} [X-\mu] \right) \tag{4}$$

where Σ_i is the 6×6 covariance matrix, and μ is the six-component mean vector. $|\Sigma_i|$ and Σ_i^{-1} are the determinant and inverse of the ith covariance matrix, respectively. During the training phase the mean vector, μ_i, and covariance matrix, Σ_i, are calculated from the training data for each class ω_i. To classify a pixel described by the feature vector X, we calculate the a posteriori probability, $P(\omega_i|X)$, that the pixel belongs to class ω_i. It is given by

$$p(\omega_i|X) = \frac{p(X|\omega_i)p(\omega_i)}{p(X)}, \quad p(X) = \sum_{i=1}^{24} p(X|\omega_i)p(\omega_i) \tag{5}$$

Its value is determined by the Bayesian decision rule, which involves the computation of the Mahalanobis distance (16).

$$d_i(X) = \ln p(\omega_i) - \frac{1}{2}|\sum_i| - \frac{1}{2}(X-\mu)' \sum_i^{-1} (X-\mu) \tag{6}$$

In the feature space, there are 24 clusters. Each pixel is assigned to the class with the nearest mean, according to this decision function.

RESULTS

It is reasonable to assume that the image normalization should result in improved accuracy of subsequent pixel classification. The performance of registration and other image normalization steps will be assessed with a Bayesian classifier and the M-FISH database (13) that we have established. Normal chromosomes were used for validation because the karyotyping of normal chromosomes is easily established with the help an expert cytogeneticist. It is expected that the improved classifier in turn would significantly identify chromosome abnormalities when applied to cancer and genetic diagnosis.

Effect of Background Flattening

To test the effect of background subtraction on the accuracy of classifier, we performed the classification of chromosomes with and without background correction. This comparison of classification accuracy is presented in Table 2. We found that one must fit the cubic background surface through points that are near to and points that are distant from the chromosomes. In addition, the background class should be trained on pixels that fall near to rather than far from the chromosomes. This decreases the misclassification of edge pixels, although distant background pixels are still classified correctly.

Effect of Feature Selection on Classification Accuracy

We performed the principal component analysis (16) to eliminate redundancy among multiple spectral channels. The dimensionality of the six-dimensional dataset is decreased by keeping only the five largest principal components of the data. In other words, we drop the least important feature. We present comparisons in Table 3. This result indicates that dimensionality reduction can improve the classification accuracy in some instances. This confirms the conclusion that five spectral channels (excluding DAPI) are usually sufficient for the discrimination of these 24 classes of chromosomes by color combination.

Table 3
Comparison of Pixel-by-Pixel Classification Rate (%) of M-FISH Image Sets With and Without Dimension Reduction

Testing image set	Without dimension reduction	With dimension reduction
1	75.53	88.63
2	71.92	84.67
3	89.40	90.35
4	90.56	85.26
5	90.25	89.34
Average	83.53	87.65

Table 4
*Comparison of Computational Time With and Without Wavelet Searching**

Similarity measure	Without wavelet	Haar wavelet	Spline wavelet
LSE	87.02	37.26	42.29
MI	105.26	43.28	58.43
X-Cor	84.23	41.51	30.00

*LSE, least square error; MI, mutual information; X-cor, cross correlation.

Evaluation of Image Registration

The rationale behind our experiment concerns (a) whether the proposed muitiresolution wavelet approach can significantly decrease the computational time and improve the accuracy of searching transformation parameters; (b) what is the best similarity metric for M-FISH registration; and (c) whether the registration can lead to improved classification accuracy. To answer these questions, we made the following comparisons.

Computational Speed With Versus Without Multiresolution Approach

The automated registration of five image channels is a time-consuming process. The registration algorithm is expected to be accelerated by multiresolution wavelet-based searching. To evaluate the computational improvement brought about by the wavelet algorithm, we tabulated the CPU time for processing an image by using the Haar wavelet and a spline wavelet filter of size 4 (16). We also listed them as a function of different similarity criteria, as in Table 4. These similarity measurements include LSE, MI, and cross correlation. The results confirm that the multiresolution approach increases computation speed. However, larger filters may lead to longer computational times.

Accuracy With Versus Without Multiresolution Approach

To test the robustness of multiresolution approach to the choices of initial values when implementing the nonlinear optimization (equation [1]), we simulated an image with known translation and rotation parameters that were used as ground truth to test the accuracy. Figure 4 shows a comparison with versus without multiresolution approach. The image was translated in the horizontal and vertical directions by two pixels and rotated by 2 degrees. With multiresolution Harr or spline wavelets, we obtain a more accurate estimations than without multiresolution.

Comparison of Different Similarity Metrics

To evaluate several similarity metrics such as the MI, LSE, and normalized cross correlation for the registration, we performed the test on the simulated images with these three different similarity measurements. Figure 5 shows the comparison among these three different metrics, where the same simulated image for comparing accuracy was used. From the experiment we concluded that MI is better than normalized cross correlation and LSE approximation methods.

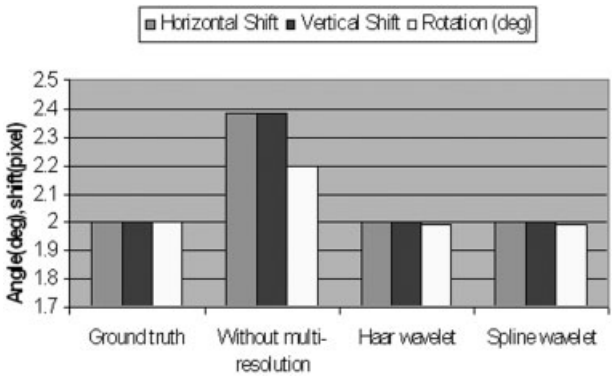

FIG. 4. Accuracy with versus without wavelet approach.

Effect of Registration on Accuracy

The classification of each pixel in the chromosome depends on the simultaneous pixel values in the six color channel images. Figure 6 shows an example of image misalignment that can be corrected with registration techniques. Obviously, proper registration will improve the accuracy of pixel classification, particularly near chromosome edges and boundaries of rearrangements. This would increase the reliability of this technique in identifying small anomalies for cancer and genetic disease diagnosis.

By using the ADIR M-FISH database (13), we conducted experiments using the proposed registration technique to evaluate the effect of registration on classification accuracy. The results of classification on several different image samples labeled with the Vysis probe set are listed in Table 5. These M-FISH images were collected with good optical component alignment. It indicates that, even after better optical component alignment, the software approach of registration for each different image samples increases the correct pixel classification rate.

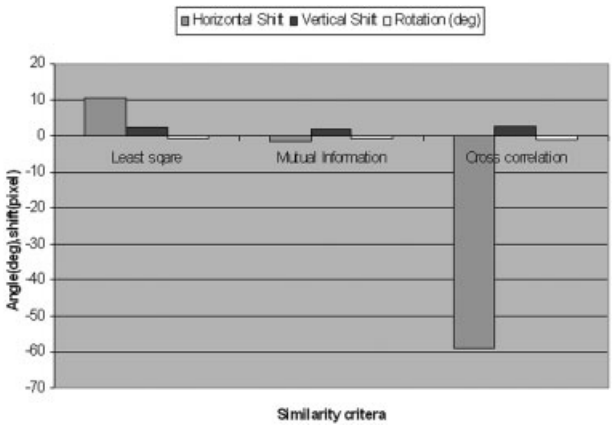

FIG. 5. Comparison of accuracy with different similarity metrics.

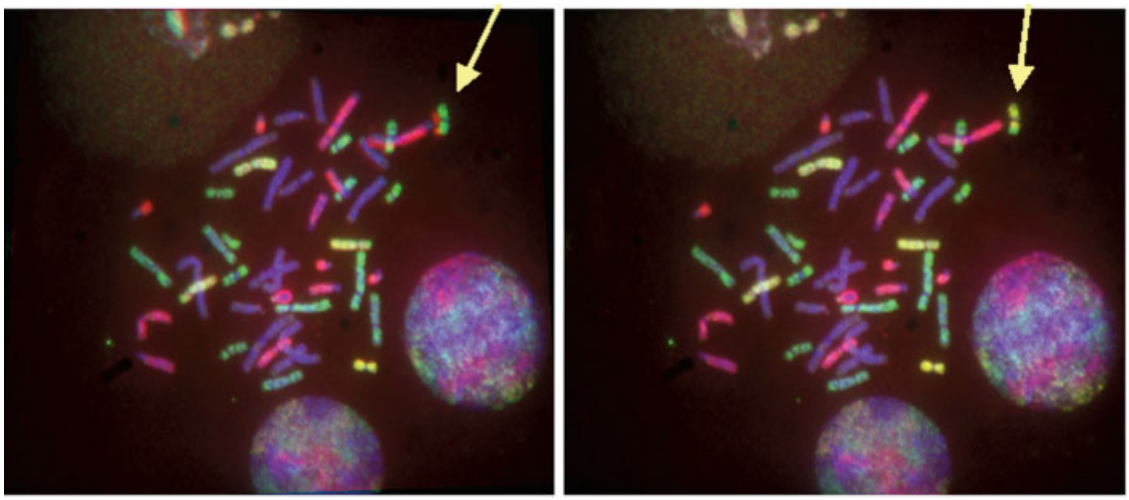

FIG. 6. Effect of registration on multichannel images. Three images from channel spectrum (s). aqua, Far Red, and DAPI are combined and displayed using red, green, and blue colors. Before registration, there is misalignment (left, arrow). After registration, good colocalization occurs (right, arrow). This misalignment, without correction, makes subsequent pixel classification less accurate. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

The increase of classification accuracy brought by registration can also be verified by other classifier. Table 6 shows the classification by a fuzzy c-means algorithm (25). In other words, regardless of the classifier used, misregistration is an important source of error that must be corrected.

DISCUSSIONS AND CONCLUSIONS

Current M-FISH instruments typically do generally accurate karyotyping but exhibit misclassifications of numerous pixel regions. The pixel misclassification errors result from a number of factors, including biochemical noise, electronic noise, spectral overlap, and misregistration. An image normalization step is required for improving accuracy in pixel classification. Improved classification accuracy will greatly decrease the size and number of misclassified pixel regions. This would allow smaller rearrangements to be identified and better enable the technique to resolve complex rearrangements when applied to patient diagnosis, thus providing greater applicability of the technique in automated karyotyping.

In this report we have introduced several image normalization approaches, especially a wavelet-based approach for multispectral FISH image registration. By performing the registration in a multiresolution framework, accuracy and speed of registration algorithm are improved. The validation of these approaches was conducted on normal chromosomes. Even for the tested images collected with a good optical component alignment, we found approximately one to three pixel shifts and 1- to 2-degree rotations. Correction of these positional errors improves classification rate. We anticipate that the registration algorithm will lead to more improved accuracy in case these optical components are not well aligned. It is reasonable to believe that the improved accuracy of classifier brought about by these normalization approaches would result in more accurate detection of subtle genetic rearrangements for cancer diagnosis and research. We plan to evaluate these approaches on patient data in collaboration with cytogeneticists in the future. In this work, we have assumed that the mulichannel images follow a rigid transformation. It is expected that a more general assumption on the nonrigid transformations between multiple channels could lead to further improvement. This may involve the use of fluorescent beads (18) and design of different computational methods. Registering the other color images to the DAPI image may be least optimal because the optical performance of the microscope system is most different in the ultraviolet/blue region where DAPI excites and emits. More appropriate would be to use a fluorochrome in the center of the visible spectrum. In addition, our future work will develop automated background subtraction by using more sophisticated spline fitting approaches (22).

The developed registration algorithm in principle can be applied to any registration problem such as protein colocalization and two-channel microarray image registra-

Table 5
Comparison of Pixel-by-Pixel Classification Rate (%) of M-FISH Image Sets With and Without Registration Using a Bayesian Classifier

Training/test image set	Without registration	With registration
V1301	46.5042	49.2528
V1306	86.5797	87.7719
V1303	68.9885	70.2981
V1304	45.2157	47.9901
V1308	52.5654	56.1503
V1309	48.4679	50.7460
Average	58.0536	60.3682

Table 6
Comparison of Pixel-by-Pixel Classification Rate (%) of M-FISH Image Sets With and Without Registration Using a Fuzz C Means Classifier

Training image	Testing image	Without registration	With registration
PNG2	A0202	83.98	84.32
PNG3	A0215	93.10	93.49
V291162	A402	85.50	86.07
V29962	A0620	86.23	86.77
A101	A403	83.97	84.32
A0325	A0218	89.54	90.38
Average		87.0533	87.5583

tion. The proposed multiresolution approach offers the advantages of improved accuracy and speed over conventional registration methods.

ACKNOWLEDGMENTS

The classification code and some of the results were prepared by M. Sampat (University of Texas, Austin). Wade Schwartzkopf prepared the ADIR M-FISH database. Some of the results were generated by Ashok Dandpat (University of Missouri). Vibeesh Bose and Hyohoon Choi also contributed to software development. We thank P. Rogan and J. Knoll at the Genetics Laboratory of the Children's Mercy Hospital for fruitful discussions.

LITERATURE CITED

1. Ried T, Baldini A, Rand TC, Ward DC. Simultaneous visualization of seven different DNA probes by in situ hybridization using combinatorial fluorescence and digital imaging microscopy. Proc Natl Acad Sci USA 1992;89:1388–1392.
2. Schröck E, du Manoir S, Veldman T, Scoell B, Wienberg J, Ferguson-Smith MA, Ning Y, Ledbetter DH, Bar-Am I, Soenksen D, Garini Y, Ried T. Multicolor spectral karyotyping of human chromosomes. Science 1996;273:494–497.
3. Speicher MR, Ballard SG, Ward DC. Karyotyping human chromosomes by combinatorial multi-fluor FISH. Nat Genet 1996;12:368–375.
4. Nederlof PM, van der Flier S, Wiegant J, Raap AK, Tanke HJ, Ploem JS, van der Ploeg M. Multiple fluorescence in-situ hybridization. Cytometry 1990;11:126–131.
5. Speicher MR, Ballard SG, Ward DC. Computer image analysis of combinatorial multi-fluor FISH. Bioimaging 1996;4:52–64.
6. Eils R, Uhrig S, Sätzler K, Saracoglu K, Bolzer A, Chassery J-M, Ganser M, Speicher MR. An optimized and fully automated system for fast and accurate identification of chromosomal rearrangements by multiplex-FISH (M-FISH). Cell Genet Cytogenet 1998;82:160–171.
7. Morrison LE. Chromosome analysis by multicolor ISH using direct-labeled fluorescent probes. Clin Chem 1993;39:733–734.
8. Morrison LE, Legator MS. Two-color ratio-coding of chromosome targets in FISH: quantitative analysis and reproducibility. Cytometry 1997;27:314–326.
9. Castleman KR, Riopka TP, Wu Q. FISH image analysis. IEEE Eng Med Biol 1996;15:67–75.
10. Saracoglu K, Brown J, Kearney L, Uhrig S, Azofeifa J, Fauth C, Speicher MR, Eils R. New concepts to improve resolution and sensitivity of molecular cytogenetic diagnostics by multicolor fluorescence in situ hybridization. 2001;Cytometry:44:7–15.
11. Lee C, Gisselsson D, Jin C, Nordgren A, Ferguson D, Blennow E, Fletcher J, Morton C. Limitations of chromosome classification by multicolor karyotyping. Am J Hum Genet 2001;68:1043–1047.
12. Castleman KR, Morrison E, Piper J, Saracoglu K, Schultz M, Speicher MR. Classification accuracy in multiple color fluorescence imaging microscopy, Cytometry 2000;41:139–147.
13. The ADIR M_FISH Image Database. Available at: http://www.adires.com/05/Project/MFISH_DB/MFISH_DB.shtml
14. Schwartzkopf W, Evans BL, Bovik AC. Minimum entropy segmentation applied to multi-spectral chromosome images. Proc IEEE Int Conf Image Process 2001;2:865–868.
15. Schwartzkopf W, Evans BL, Bovik AC. Entropy estimation for segmentation of multi-spectral chromosome images. In: Proceeding of the IEEE Southwest Symposium on Image Analysis and Interpretation; Santa Fe, New Mexico; 7–9 April 2002. 234–238.
16. Castleman KR. Digital image processing. Saddle River, NJ: Prentice-Hall; 1995.
17. Castleman KR. Color compensation for digitized FISH images. Bioimaging 1993;1:159–165.
18. Kozubek M, Matula P. An efficient algorithm for measurement and correction of chromatic aberrations in florescence microscopy. J Microsc 2000;200:206–217.
19. Brown LG. A survey of image registration techniques. ACM Comput Surv 1992;24:325–376.
20. Dennis J Jr, Schnabel R. Numerical methods for unconstrained optimization and nonlinear equations. SIAM, Classics in Applied Mathematics 16, 1996.
21. Press WH, Teukolsky SA, Vetterling WT, Flannery BP. Numerical recipes in C: the art of scientific computing. New York: Cambridge University Press; 1988.
22. Unser M. splines: a perfect fit for signal and image processing. IEEE Signal Process Mag 1999;16:22–38.
23. Thevenaz P, Ruttimann U, Unser M. A pyramid approach to subpixel registration based on intensity. IEEE Trans Image Process 1998;7:27–41.
24. Sampat MP, Bovik AC, Aggarwal JK, Castleman KR. Supervised parametric and non-parametric classification of M-FISH chromosome images. Pattern Recogn (in press).
25. Wang YP, Danpat A, Castleman KR. Classification of M-FISH images using fuzzy C-means clustering algorithm and normalization approaches. Asilomar Conference on Signals, Systems, and Computers; 7–10 November 2004.
26. Wang Y-P, Wu Q, Castleman KR, Xiong Z. Chromosome image enhancement using multiscale differential operators. IEEE Trans Med Imaging 2003;22:685–693.