Classification of Subject Motion for Improved Reconstruction of Dynamic Magnetic Resonance Imaging

Feiyu Chen

Department of Electrical Engineering

ABSTRACT

Subject motion is a significant problem in magnetic resonance imaging (MRI). In this project, unsupervised learning using K-means clustering, and supervised learning using logistic regression (LR) and support vector machine (SVM) are implemented and applied to solving (1) classification of MR images with/without motion, and (2) estimation of current motion state based on previous low frequency navigator signals. 105 MR images with labeling of motion/no-motion and 40000 navigator signals are collected as the training and testing data. Over 10 features were designed and manually selected for each learning approach based on principle component analysis (PCA). Results show that, for unsupervised learning, the K-means algorithm achieves an accuracy of ~83% for classification of MR images and a root-mean-squared error (RMSE) of 5.3% in prediction based on navigator signals. For supervised learning, SVM using Gaussian kernels achieves the best accuracy of ~90% for classification of non-motion and motion images, and a RMSE of ~1.7% for estimation of motion states based on navigator signals. These approaches provide a possibility for advanced reconstruction algorithms to perform more accurate motion corrections with or without a large training set, and may improve the quality of clinical images.

INTRODUCTION

Among all the medical imaging modalities, MRI is one of the best candidates for evaluation and diagnosis of diseases, because of its high spatial-resolution, good soft-tissue contrast and no ionizing radiation. However, MRI requires relatively long scan times, and therefore is more sensitive to subject motion during exams. Subject motion can create imaging artifacts such as imaging blurring, ghosting, and reduced SNR, and degrade image quality. This limits its wide use in clinical diagnoses. In order to overcome this disadvantage, several approaches have been developed for improving speed and reducing motion artifacts during reconstructions of acquired MRI signals. However, these approaches require classification and estimation of motion, which is usually difficult to achieve.

The fundamental goal of this project is to classify the images with/without motion based on DICOM images. The second objective of this project is to estimate the current motion state based on previous information on patient positions. With the help of unsupervised and supervised learning models, hopefully MRI scanners can automatically determine the significance of motion artifacts, and perform appropriate corrections and reconstructions based on prior information of motion.

RELATED WORK

Both unsupervised learning and supervised learning have been applied to solving MRI problems. However, only unsupervised learning has been used to classify MR images for cardiac perfusion applications[1]. Few attempts have been made to classify abdominal images and mixed images with different body parts. Supervised learning approaches, such as support vector machine (SVM), have been used to classify functional MRI data of brain states[2][3]. However, it has not been applied to prediction of motion states and subject positions, which is the main focus of this project.

DATASET & FEATURES

1. Motion artifacts in MRI

A good understanding of the motion artifacts in MRI is helpful for designing features for learning. As data are directly acquired in the frequency domain in MRI, motion artifacts of the displayed images are different from common motion artifacts in optical photos. According to the Fourier Transform theories, subject motion will cause phase shift in the frequency domain. Assuming several parts of the data in frequency domain are corrupted by motion, the image will show motion artifacts as blurring and ghosting, which contain shifted replicas of the original image without motion, as shown in Fig.1.

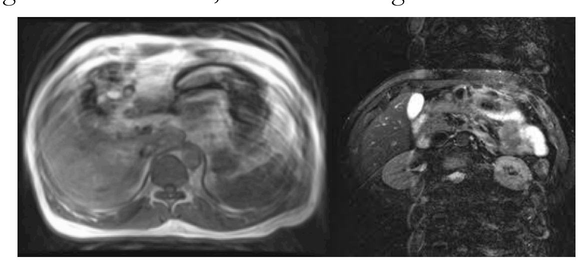

Fig. 1 Illustration of motion artifacts in abdominal MRI

2. Low frequency navigator signals for motion

Low frequency signals are acquired before and after each image frame as navigators for the real-time motion state, as the distance of subject motion can be estimated directly from the phase of the low frequency signal (Fig. 2). However, in contrast-enhanced dynamic MRI, many other factors contribute to the motion estimation, among which contrast change is the most significant one. This leads to wrong estimation of current motion state. In this project, this problem would be solved based on previous noncontaminated signals, which would be beneficial for clinical imaging.

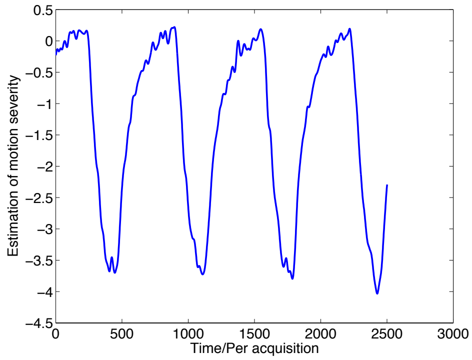

Fig. 2 Illustration of low frequency navigator signals in MRI

3. Samples of images and navigators

For the first objective of the project, 57 images with motion and 48 images without motion were collected in experiments conducted at Stanford, and online searches of public MR images. Two sets are formed based on these images. The first set contains images with similar regions of the abdomen. The second set contains images with different body parts, including the brain, the liver, the pelvis, and the heart. For the second objective, 40000 navigator signals are acquired in a continuous scan with free breathing, where one breathing period is ~500 time points.

4. Pre-processing and features

4.1 Images

Pre-processing of images includes the following steps:

(a) Resize images into the same size of 256*256;
(b) Divide each image equally into blocks;
(c) Compute the features for each block.

The features being tested include:

(1) The mean value of all the pixels in a block;

(2) The rank of the image, estimated from the singular value decomposition (SVD);

(3) The standard deviation of pixels in a block;

(4) The number of points on edges;

(5) The sum of intensity for the 25% center of images;

(6) The sum of intensity for the edges of the image;

(7) The ratio between 5 and 6;

(8) The total variation of the image;

(9) The ratio between the center intensity and edge intensity of the frequency domain;

(10) The top 10% eigenvalues of the image;

(11) Phase of the Fourier transform of the image.

4.2 Low frequency navigator signals

Before extracting features of navigator signals, they are first discretized into 10, 20, or 40 states. The previous 500 time points before the current point are extracted in a sliding-window pattern.

The features being tested include:

(1) A vector of previous signal values;

(2) The range of previous signal values;

(3) The mean of previous signal values;

(4) The slope from the beginning to the end;

(5) The center value of previous signals;

(6) The coefficients of 2D polynomial fitting;

(7) The index of points with max and min values;

(8) The beginning slope and ending slope;

(9) The coefficients of linear fitting.

4.3 Normalizations and principle components analysis (PCA)

After extracting the features, the values of each feature are normalized with the statistical average μ and standard deviation σ with $x' = \dfrac{x-\mu}{\sigma}$. PCA is subsequently performed on all training data to project the features onto a subspace with maximized differences and efficiency. A threshold of 5% energy of the first principle component is used to determine the new feature space with reduced dimension.

METHODS

1. Learning models

1.1 Unsupervised learning using K-means clustering

The unsupervised learning model being implemented is the K-means method. This method uses the geometric distance between samples as the metric for classify and cluster a set of samples in an iterative way. For classification of images with motion and without motion, the number of clusters is set to be two. For estimations of the severity of motion, the number of clusters is set to the number of discretized states.

1.2 Supervised learning using LR and SVM

LR uses the sigmoid function as the regression model to fit the labels with input features. It is implemented in MATLAB using the stochastic gradient ascent algorithm with coefficient α =0.005. SVM uses support vectors to determine the boundary between classes with different labels. The kernels in the dual optimization problem of SVM determine the hyperspace of features. Linear kernels (SVML) and Gaussian kernels (SVMG) are two common kernels and are implemented in this project. The LIBLINEAR and LIBSVM toolbox[4] are used as implementations of SVML and SVMG, respectively.

2. Evaluations

2.1 Evaluation of classifications based on MR images

Evaluations of the learning theories are conducted based on the true positive (TP), true negative (TN), false positive (FP), and false negative (FN) counts in leave-one-out crossing-validation experiments, as the size of training data is relatively small. The accuracy, specificity, and sensitivity of the learning theory are further computed based on these indices.

2.2 Evaluation of classifications based on low-frequency navigator signals

Since the number of states is much greater than two, using true and false cannot comprehensively show the accuracy of each learning model. Therefore, the root-mean-squared error (RMSE) is calculated between the estimates and the reference labels in two-fold cross-validation experiments, due to the large size of training data. The RMSE is normalized with the total number of states to provide a fair comparison between different discretization of motion states.

$$RMSE_{Normalized} = \frac{\sqrt{\sum (S_{est} - S_{true})^2}}{\sqrt{N_{Test} \cdot N_{States}}}$$

where S means the signal, N_{test} is the length of test data, and N_{states} is the number of discretized states.

RESULTS & DISCUSSION

1. Classification of images with and without motion

The accuracy of each learning approach is evaluated for (1) a set of abdominal images, and (2) a set of images of different body parts. All learning approaches achieve similar performance for the first dataset. Compared with unsupervised learning using K-means, supervised learning approaches generally achieve higher accuracy of ~90% for the images of different body parts.

1.1 Unsupervised learning using K-means

The K-means approach reaches an accuracy of 83% for both cases, as shown in Table 1. Compared with the first dataset, K-means for the second dataset has higher sensitivity (100%) but lower specificity (63%).

Table 1 Unsupervised learning for classification of MR images with motion and without motion

Images	Dataset 1: Abdominal images only	Dataset 2: images with variety of regions
Features	[1 2 4 8 9]	[1 2 3 4 8 9]
#Blocks	[3 1 4 3 1]	[1 1 1 1 1 1]
TP	29	57
TN	34	30
FP	4	18
FN	9	0
Sensitivity	76%	100%
Specificity	89%	63%
Accuracy	83%	83%

1.2 Supervised learning using LR and SVM

All the three supervised learning approaches perform well for the first dataset. SVM with Gaussian kernels achieves slightly higher accuracy (2% higher). For the second dataset, the three approaches achieve similar accuracy of ~90%. However, SVML achieves higher specificity (94%), while SVMG achieves higher sensitivity (98%).

Table 2 Supervised learning for classification of MR images with motion and without motion

Images	Dataset 1: Abdominal images only			Dataset 2: images with variety of regions		
Features	[1 2 4 8 9]			[1 2 4 6 7 9 11]		
#Blocks	[3 1 4 3 1]			[4 1 2 4 1 1 1]		
Models	LR	SVML	SVMG	LR	SVML	SVMG
TP	37	37	38	50	50	56
TN	38	38	38	43	45	39
FP	0	0	0	5	3	9
FN	1	1	0	7	7	1
Sensitivity	97%	97%	100%	88%	88%	98%
Specificity	100%	100%	100%	90%	94%	81%
Accuracy	98%	98%	100%	89%	90%	90%

1.3 Analysis of accuracy

1.3.1 Errors in unsupervised learning with K-means

The main error of K-means for classification of the second dataset comes from mislabeling of points on the boundary of the two classes, as shown in the red ellipsoid in Fig. 3. The scatter plot of correct labeling also reveals it is difficult to separate the two image classes using only these two features, because of the mixture of labels in the center.

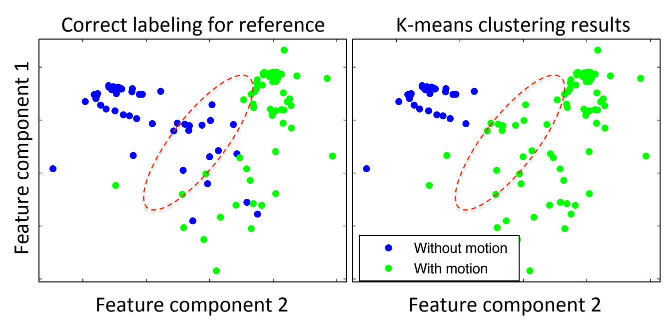

Fig. 3 Illustration of K-means clustering for MR images with/without motion in the second dataset. The x and y axes denote the two major feature components

1.3.2 Analysis of #blocks for supervised learning

The size/total number of blocks influences the accuracy of all supervised learning methods. As shown in Fig. 4, both SVML and LR show fluctuations of accuracy with changes of the number of blocks, while the accuracy of SVMG remains stable for different number of blocks. However, all the three approaches show a trend of accuracy increase with increasing number of blocks. A possible reason is, the efficiency of information increases with increased number of blocks, which provides more features for training and classification.

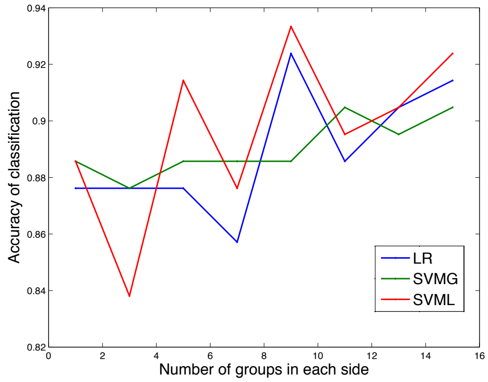

Fig. 4 The accuracy of classification varies with changes in number of blocks in each side

1.3.3 Training set size for supervised learning

Training set size is another key factor for the accuracy of supervised learning. As shown in Fig. 5, all the three learning approaches demonstrate a tendency of accuracy increase with increased training set size. When the training set size is less than 40, there are fluctuations in the classification accuracy for all three methods. Therefore, it is necessary to maintain a training set size of over 40.

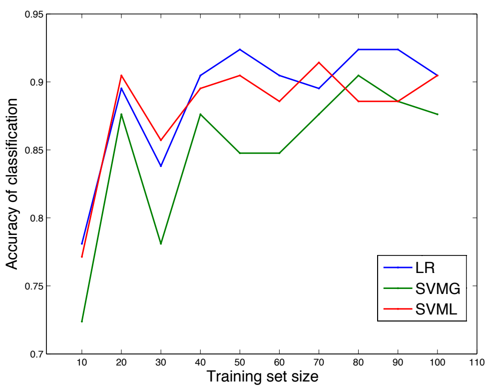

Fig. 5 The accuracy of classification varies with changes of the training set size

2. Estimation of the position of motion based on low-frequency navigator signals

2.1 Unsupervised learning using K-means

As there are multiple classes in this classification problem, and the differences between difference classes are relatively small, unsupervised learning may encounter some difficulty in generating appropriate clusters. However, with appropriate features, the K-means approach achieves a RMSE of less than 10% for 20000 testing data points with feature set [1 7 8 9], as shown in Table 3. The curve in Fig.5 also shows K-means can follow the signal evolutions well, despite some minor delays in the decreasing periods.

Table 3 RMSE for estimation of motion states based on navigator signals and 20000 training data

#States	Features	K-means	LR	SVML	SVMG
40	[1 7 8 9]	5.26%	3.51%	6.74%	1.67%
20	[1 7 8 9]	8.02%	3.08%	6.34%	1.22%
10	[1 7 8 9]	10.07%	3.21%	4.00%	1.58%

2.2 Supervised learning using LR and SVM

Supervised learning approaches achieve higher accuracy of less than ~6% compared with the K-means approach. SVMG achieves the lowest RMSE among all the three supervised learning approaches, while LR performs better than SVML (Table 3). The signal evolution curves also show SVMG fits the curve better than LR. LR has higher accuracy and less noise than SVML at the peaks and troughs of the curve (Fig. 5).

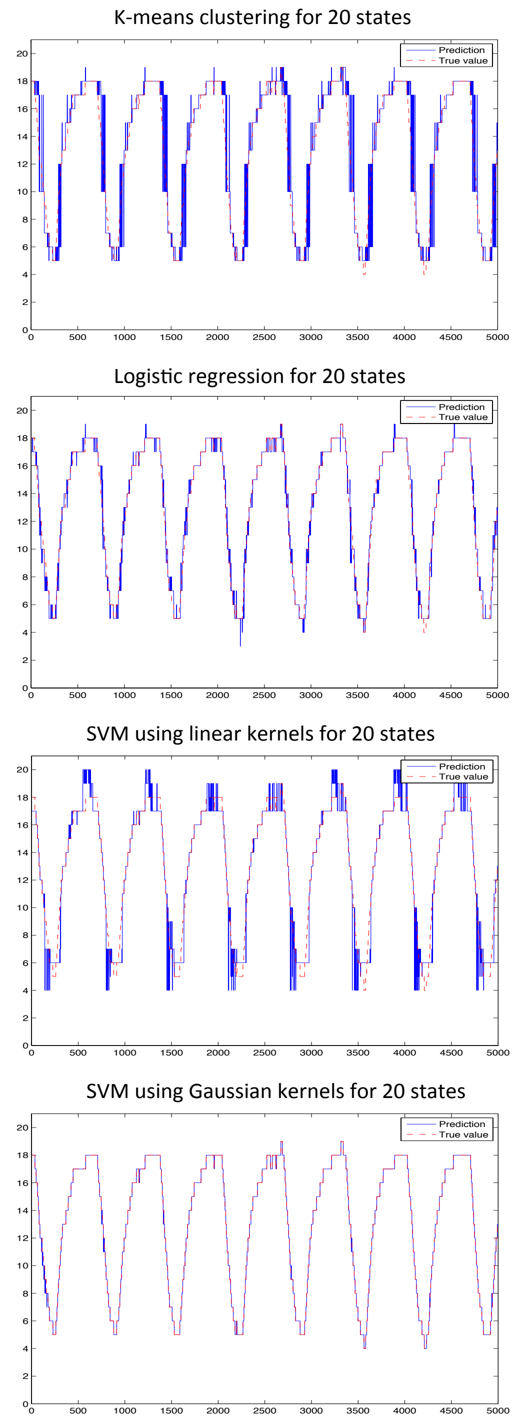

Fig. 5 Illustration of estimation based on low frequency navigator signals in MRI using K-means, LR, SVML, and SVMG for signals with 20 discretized states

2.3 Analysis of performance

2.3.1 Analysis of the size of prior information

The size of prior information, i.e., the length of time series being used to generate features, may have an effect on the accuracy of learning models. In specific cases with periodic motion, the period of breathing is ~500 time points, which makes the choice more important. As shown in Fig. 6, only LR shows an improvement of accuracy (RMSE decrease) with increasing length of prior information. K-means and SVMG remains insensitive to the size of previous points, and SVML shows significant fluctuations with this change.

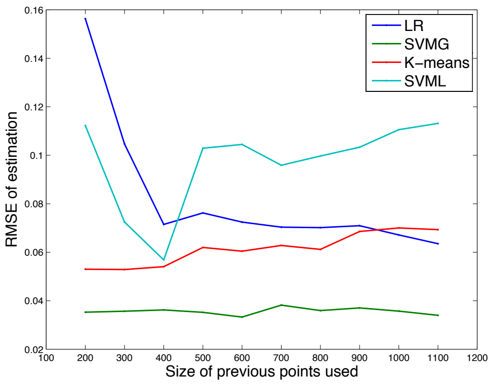

Fig. 6 Changes of RMSE with increasing size of prior information (#points used for feature extractions) using a training set size of 5000 for estimation of motion states based on navigator signals

2.3.2 Analysis of training set size for supervised learning

The relationship between the RMSE of testing data and training size is shown in Fig. 7. As the training size increases, the RMSE of LR, SVML, and SVMG for supervised learning decrease. Among the three supervised learning methods, SVMG preserves the highest accuracy even with few training data.

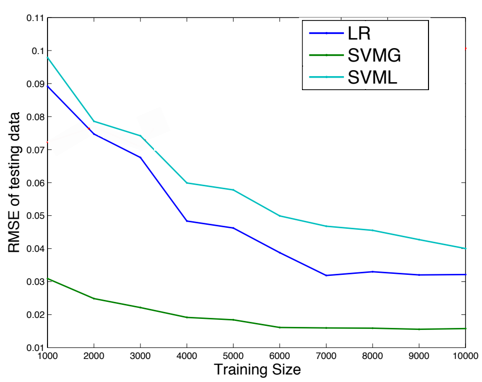

Fig. 7 RMSE of the testing data with increasing training size using 500 points of previous signals for estimation of motion states based on navigator signals

2.3.3 Number of discretized states

According to Table 3, for unsupervised learning with K-Means, the estimation accuracy increases with the number of discretized states. However, for all three supervised learning approaches, the estimation accuracy remains similar or fluctuates slightly as the number of states increases. This shows the number of discrete states has little impact on the estimation accuracy.

CONCLUSION & FUTURE WORK

1. Choice of learning models for motion classification and estimation

If a large amount of training data with correct labeling is available, supervised learning should be used for both classification of images with/without motion and estimation of motion states based on navigator signals. Among the three supervised learning approaches, SVM with Gaussian kernels provides the most stable and accurate labeling for both cases. LR has slightly worse performance than SVMG, but it's also easier to understand and implement.

If a reasonable amount of training data cannot be achieved, the K-means approach can give acceptable classification of images with/without motion, and good estimation of motion states. While its accuracy is not as good as supervised learning algorithms, it is not susceptible to the length of time points used to generate features when doing estimation based on navigator signals.

2. Limitations of current approaches

First, the size of training sets for classification of images with/without motion is limited, since no big public database is available online. This limits the performance of all the three supervised learning approaches (Fig. 5). Second, there are many kinds of subject motion during clinical studies. Therefore, a multi-level classification may be useful to improve the accuracy. Third, due to hardware limitations, the motion signal can only be discretized to 40 states. More states would be beneficial to subsequent motion corrections and reconstructions.

3. Future work

Future work can be done in (1) exploring more advanced learning models, (2) collecting more samples of images with/without motion, and (3) conducting motion correction based on the results of classification and estimation of motion to see how machine learning can help with motion corrections and reconstructions of MR images.

REFERENCES

[1] Stegmann, Mikkel B., Hildur Ólafsdóttir, and Henrik BW Larsson. "Unsupervised motion-compensation of multi-slice cardiac perfusion MRI." Medical Image Analysis 9.4 (2005): 394-410.

[2] Mourão-Miranda, Janaina, et al. "Classifying brain states and determining the discriminating activation patterns: support vector machine on functional MRI data." NeuroImage 28.4 (2005): 980-995.

[3] Wang, Ze, et al. "Assessment of functional development in normal infant brain using arterial spin labeled perfusion MRI." Neuroimage 39.3 (2008): 973-978.

[4] Chih-Chung Chang and Chih-Jen Lin, LIBSVM : a library for support vector machines. ACM Transactions on Intelligent Systems and Technology, 2:27:1--27:27, 2011. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm