Compressive Sensing for Gait Recognition

Sabesan Sivapalan, Rajib Kumar Rana*, Daniel Chen, Sridha Sridharan, Simon Denmon and Clinton Fookes
Image and Video Research Laboratory,
Queensland University of Technology,
GPO Box 2434, 2 George St.
Brisbane, Queensland 4001.
{*sivapalen.sabesan*, *daniel.chen*, *s.sridharan*, *s.denman*, *c.fookes*}@qut.edu.au, *rajeevrans@gmail.com*
*Information & Communication Technologies Centre, Commonwealth Scientific and Industrial Research Organisation, Australia.

***Abstract*—Compressive Sensing (CS) is a popular signal processing technique, that can exactly reconstruct a signal given a small number of random projections of the original signal, provided that the signal is sufficiently sparse. We demonstrate the applicability of CS in the field of gait recognition as a very effective dimensionality reduction technique, using the gait energy image (GEI) as the feature extraction process. We compare the CS based approach to the principal component analysis (PCA) and show that the proposed method outperforms this baseline, particularly under situations where there are appearance changes in the subject. Applying CS to the gait features also avoids the need to train the models, by using a generalised random projection.**

I. Introduction

Compared to other biometrics, gait has attracted significant attention in recent years because of its unique advantages, which other biometrics may not offer. Mainly, it can be used with low-resolution video feeds, and be acquired at a distance without alerting the subject.

There are a number of methods being investigated to extract gait features, generally classified as appearance-based or model-based. Model-based techniques gather gait dynamics directly by modelling the underlying kinematics of human motion, where as appearance based methods try to establish correspondence between successive frames based upon the implicit notion of what is being observed [1]. Both approaches attempt to address the main challenges in gait recognition such as invariance to changes in view, surface, clothes etc.

Many appearance based techniques, such as gait energy images (GEI) [2], uses an image as its features and thus are of a high dimensionality. Principal component analysis (PCA), linear discriminant analysis (LDA) and independent component analysis (ICA) are commonly used to reduce the dimensionality in gait recognition algorithms. These techniques learn from a training set and therefore suffer from limited generalisation.

In this paper, we evaluate compressive sensing (CS) to replace the traditional dimensionality reduction techniques used in gait recognition approaches, using GEI as the base feature extraction technique.

Compressive sensing [3], [4] is an emerging field of information theory. In the last decade, compressive sensing theory has been widely explored by the image processing research community and has been used in many applications such as human face recognition [5], [6], facial expression recognition [7], [8] and human action recognition [9]. Our proposed method is motivated by one of the pioneering face recognition works reported in [5].

The reason for choosing compressive sensing is twofold. First, we believe in the hypothesis that random features can be as good as sophisticated features provided the number of random features is sufficiently large ([5] shows that compressive sensing can be efficiently used to recognize face image using random features). Second, the compressive sensing based recognition approach does not require any training process, which makes it compelling in the domain of biometric identification.

In our paper, we use the gait energy image, introduced by Han and Bhanu [2] for extracting gait features. Gait energy image based feature extraction is a popular appearance based method, since GEI represents gait using single image without loosing temporal information. By averaging frames in the gait sequence, segmentation noise in individual frames can be overcome. Many extensions to the initial GEI technique have been proposed, including the enhanced gait energy image (EGEI) [10] and the shifted energy image (SEI) [11]. However, most of them try to address specific limitations such as changes in clothing style etc, and performance of these algorithms does not significantly improve upon GEI. Therefore, we have chosen GEI to evaluate the proposed method.

In our proposed approach, first a dictionary is formed using the gallery of GEIs of different subjects from different classes (normal walking, with a bag, etc.). Intuitively, a test GEI belonging to one of the classes of the dictionary can be sparsely represented as a linear combination of GEIs of that particular class. If a test GEI does not belong to any of the classes in the dictionary, it will be represented as linear combination of multiple class subjects, i.e. the representations will not be sparse. Sparse representation of a test image in terms of a dictionary is non-trivial. We employ compressive sensing to compute this sparse representation.

Generally speaking, compressive sensing allows us to reduce the dimension of a signal by taking a small number of projections and reconstructing the higher dimensional signal from the lower dimensional projections. Reduction of dimension is necessary for non-exhaustive processing. To maximize the dimension reduction, the signal under consideration needs to be sparse or compressible. A signal $f \in R^n$ is called K-sparse if it has only K nonzero elements [3]. On the other hand a signal with a small number of significant (large magnitude) coefficients and a large number of insignificant coefficients (very small magnitude close to zero) is called a compressible signal.

The GEIs in their original form are not sufficiently compressible. Therefore, we also seek to find an appropriate sparsifying basis that can sparsely represent a GEI. We then construct the dictionary using the sparse GEI and follow the prior mentioned process for recognition.

To this end we summarise our main contributions as follows:

1) We propose a novel compressive sensing approach for accurate person recognition using GEIs. To the best of our knowledge we are the first to propose compressive sensing for gait recognition using GEI.
2) Through evaluation, we demonstrate that our proposed method out performs the existing PCA at the same feature dimensionality without any training.

The paper is organised as follows. In the next section (Section II) we describe the compressive sensing framework for GEI based person recognition. Section III outlines the feature extraction using GEI. In Section IV, we summarise the experimental results and we conclude the paper in Section V.

II. Unified Compressive Sensing and Sparse Coding for Gait Based Person Identification

In this section, we describe the mechanics of the compressive sensing approach for gait recognition using GEI features. Let us consider that we have $c = 1, 2, .., C$ classes of GEIs, and there are N GEIs which are represented by N vectors $\vec{v}_1, .., \vec{v}_N \in \mathbb{R}^n$. Let us construct a dictionary by packing the vectors $\vec{v}_i, \forall_{i=1,..,N}$ in the column of a matrix $\xi \in \mathbb{R}^{n \times N}$. Intuitively, a test sample $\vec{\gamma} \in \mathbb{R}^n$ of class i can be represented in terms of the dictionary as follows,

$$\vec{\gamma} = \xi \vec{\alpha}, \tag{1}$$

where $\vec{\alpha} = [0, 0, \pi_{i1}, .., \pi_{ik}, 0, 0]^T$, π_{ij} are scalars and k is the number of GEIs per class. Clearly, solving (1) (i.e. $\vec{\alpha}$) would recognize the test image class, however, we have to identify a method to compute $\vec{\alpha}$.

A general method to find the sparse solution of (1) is to solve the following optimization problem:

$$\arg\min \hat{\alpha} = ||\vec{\alpha}||_0, \quad s.t. \vec{\gamma} = \xi \vec{\alpha}, \tag{2}$$

where $||.||_0$ denotes the ℓ_0 norm, which returns the nonzero elements of α. Note that if $n > N$, the system is over-determined. In this case, (2) can be solved in polynomial time. Typically the feature size (n) is quite high, therefore it is not practical to solve (2). For this reason, the dimensionality of n is reduced to $d << n$ (by multiplying a random projection matrix $\Phi \in \mathbb{R}^{d \times n}$ with ξ), which transforms (2) into an under-determined problem. Finding a sparse solution to an under-determined system using (2) is NP-hard. Encouragingly,

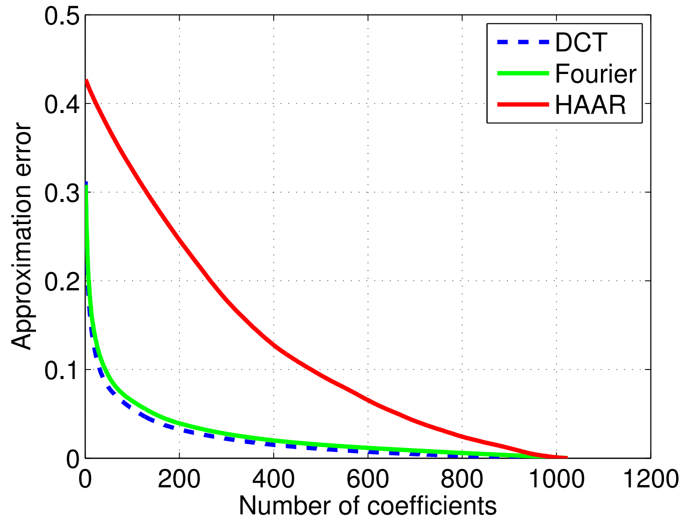

Fig. 1. Representation of GEI in different sparsifying basis. A GEI is divided into bins of size 1024. For each bin transformed coefficients are computed and approximation error for retaining different numbers of coefficients are computed. The average result for all the bins are shown in this plot.

compressive sensing shows that if $\vec{\alpha}$ is sufficiently sparse, this under-determined system can be solved using the following ℓ_1 norm minimisation problem, which will produce a similar solution to solving the ℓ_0 norm.

$$\arg\min \hat{\alpha} = ||\vec{\alpha}||_1, \quad s.t. \vec{\gamma} = \xi \vec{\alpha} \tag{3}$$

However, a sparse $\vec{\alpha}$ cannot always guarantee a unique solution to (3). Compressive sensing theory shows that if $\Theta = \Phi\xi$ obeys the restricted isometry property (RIP) [12], the under-determined system of (1) can be solved by solving (3). More encouragingly, compressive sensing also suggests a Φ that can achieve the RIP. As such, we use a Φ which is populated by sampling from a normal distribution with zero mean and variance $\frac{1}{d}$. Compressive sensing has shown that $\Phi \sim \mathcal{N}(0, \frac{1}{d})$ obeys RIP when $d >= K \log \frac{n}{K}$, where K is the measure of sparsity. Clearly, when K is small, higher reductions in the dimensionality can be be achieved.

The GEI in its original form is not sufficiently sparse i.e., K is not a small quantity. In order to make K small we sparsify the GEIs using a sparsifying basis such as wavelets, discrete cosine transform (DCT) or Fourier bases. In Figure 1, we show the approximation error for the sparse GEI under different bases. We observe that the DCT and Fourier bases have the quick decays (with DCT being slightly faster), implying that they can provide the most accurate sparse representation of GEIs.

Note that a sparsifying basis cannot be chosen only based on the guarantee of maximum sparse representation. Compressive sensing provides guidelines to choose an appropriate sparsifying basis. The sparsifying basis needs to be incoherent with the projection basis. Coherence is the measure of the highest correlation between any two elements of the projection matrix and sparsifying basis. Coherence (μ) is given by the following formula [13]:

$$\mu(\Phi,\Psi)=\sqrt{n}\max_{1\leq o,q\leq n}|\langle\phi_o,\psi_q\rangle| \tag{4}$$

Let us denote Ψ as our sparsifying basis. Generally speaking, if Φ and Ψ have many correlated elements, coherence is high, otherwise it is low. Compressive sensing requires a Ψ that has low coherence with Φ and the value, μ, within $[1,\sqrt{n}]$. After analysing the μ for different sparsifying basis, we observe that DCT has the lowest coherence with our projection matrix, which is sampled from $\mathcal{N}(0,\frac{1}{n})$. In particular, DCT and HAAR wavelets both have similar coherence with our Φ, but since DCT offers a sparser representation of GEI, we use DCT in our experiments.

Using the selected basis, we transform each of the GEIs $\vec{v_i}\forall_{i=1..n}$ to a sparse domain, using $\vec{\omega}_i=\Psi\backslash\vec{v_i}$. We then construct the dictionary $\xi=[\vec{\omega}_1\vec{\omega}_2\vec{\omega}_n]$ by packing the vectors $\vec{\omega_i}$ in columns. Then, we solve the following optimization problem to compute β:

$$\arg\min\hat{\beta}=||\vec{\beta}||_1,\quad s.t.\vec{\gamma}=\xi\vec{\beta} \tag{5}$$

As described earlier, the solution to (5) will provide $\vec{\beta}=[0,0,\pi_{i1},..,\pi_{ik},0,0]^T$, where π_i will represent the class the test GEI belongs to.

III. Gait Feature Extraction

To evaluate our algorithm, we use the GEI based gait feature extraction algorithm proposed by Han and Bhanu [2]. GEI represents the gait features in a silhouette sequence of a person over a gait cycle as a single image.

To extract out silhouettes, we use a graph cut based segmentation algorithm similar to that by Chen *et al*. [14], except that background subtraction is used to generate the foreground and background probabilities instead of a motion detection algorithm.

Gait cycles are first segmented by detecting the peaks in the width of the lower 20% of the silhouette. The silhouettes are normalised and aligned. The GEI of the k^{th} gait cycle is then computed as follows,

$$GEI(k)=\sum_{t=1}^{t=N}I_t, \tag{6}$$

where, I_t is the pre-processed silhouette of the t^{th} frame and N is the number of frames in the k^{th} gait cycle. Computed GEI images are shown in Figure 2.

Walking gait is mostly represented by the lower part of body (legs). In addition, the lower half is less susceptible to appearance changes due to situations such as the carrying of goods and hand movements. As a result, only the lower 40% of the GEI will be used as part of this paper.

IV. Experimental Results

The CASIA Dataset B [15] is used to evaluate all the experiments presented in this paper. The database consists of 124 subjects under three test classes captured from 11 cameras simultaneously. Three classes are normal walk (**nw**),

Fig. 2. Extracted silhouettes of different classes (top to bottom: **nw**,**cl** and **bg**) and computed GEIs (right most ones).

Fig. 3. Example images from CASIA dataset B. The left image shows the **nw** class, middle image is **cl** and the right image is **bg**.

walking with bag (**bg**), and walking with coat (**cl**). There are 6 sequences for **nw** and 2 sequences for both **cl** and **bg**. We use the side ($90\,\mathrm{deg}$) view to generate the GEI for all subjects and test cases. Each sequence covers at least one to two gait cycles, though only one is extracted, resulting in a total of 10 cycles for each subject. Figure 3 shows example images from the database.

Both intra and inter-class test cases are considered and receiver operating curves (ROC) are used to compare the results. In the intra-class case, we split the number of cycles in half, with the gait cycles in the first half used as the gallery and the second half used as the probe. We only consider **nw** for intra-class tests as there is insufficient data for the other classes.

The choice of dimensionality for the sparse GEI is important since it will effect the approximation error as shown in Figure 1. We evaluate this by investigating the recognition performance using a different number of components. Figure 4 plots the recognition results of intra-class tests by taking a different number of sparse components from the given 1024 components. It can be seen the optimum improvement in the results can be obtained in the range of $200-300$ and there is no significant improvement after that. This shows that the approximation error is effective for choosing the number of dimension to consider.

Figures 5 and 6 compare the results of the proposed method with PCA based GEI using the same number of dimensions (here we have chosen the first 300 components). It can be

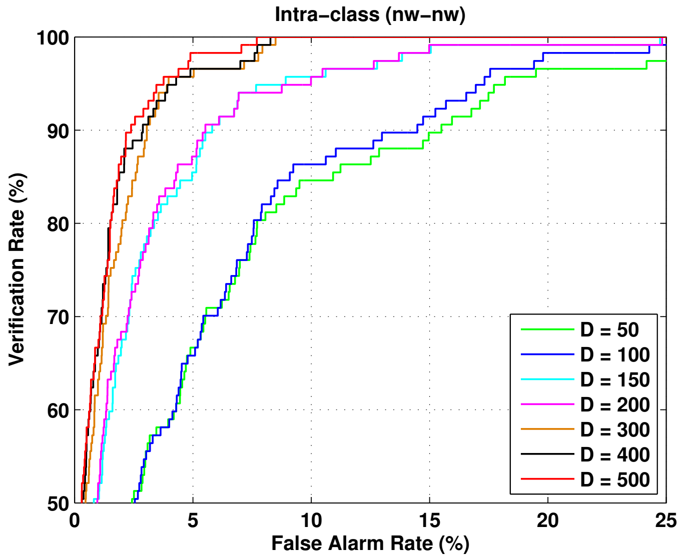

Fig. 4. Intra-class test ROC curves of proposed method using different number of sparse components.

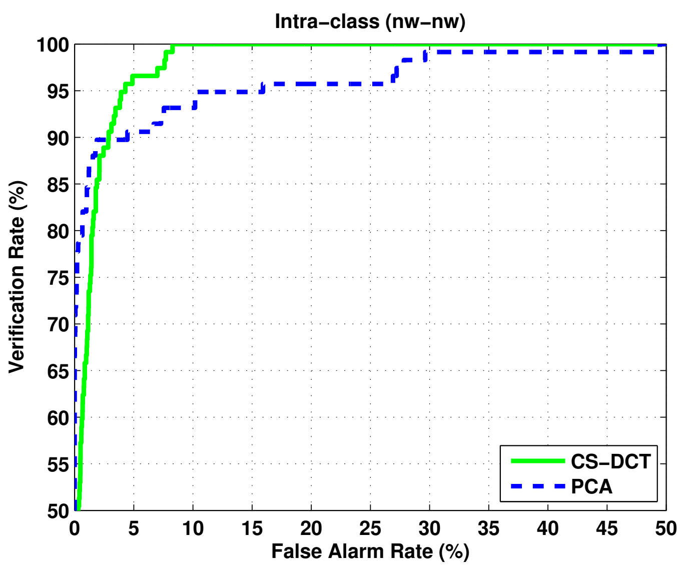

Fig. 5. ROC curve for intra-class test on the CASIA database.

seen that for intra-class tests within **nw**, the proposed method's recognition rate is 100% at a false alarm rate of 10%, 5% higher than PCA (see Figure 5). Figure 6 compares the inter-class performance between PCA and the proposed method, and it can be seen that the proposed approach improves the recognition performance significantly, and outperforms PCA by an average of more than 15% at FAR 10%. This suggests that CS is able to better tolerate the appearance variations that exist between classes.

V. Conclusion

We have demonstrated that a Compressive Sensing based dimensionality reduction technique works well for gait recognition while removing the training requirement through random generalisation basis. We have evaluated our proposed method

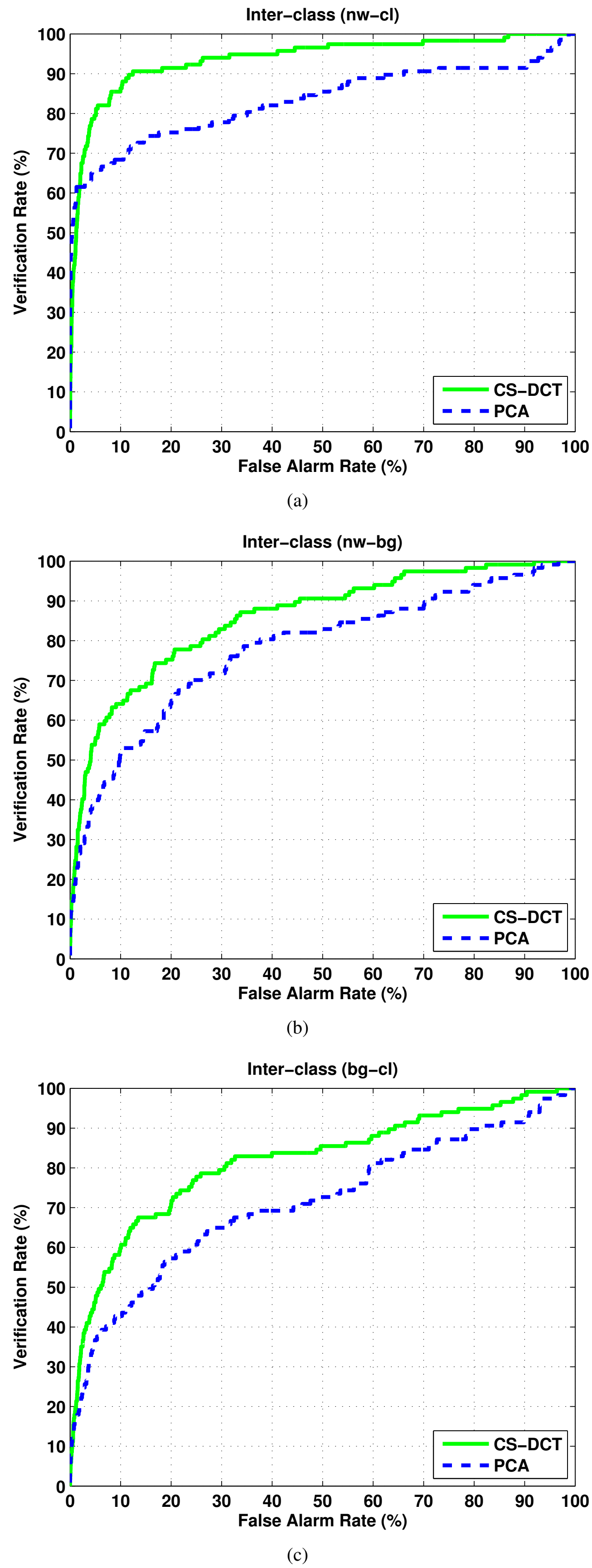

Fig. 6. ROC curves for inter-class tests on the CASIA database.

in the popular GEI based feature extraction algorithm, however other techniques could be applied using the same framework. In future, the performance of the proposed method can also be improved by learning an optimal sparsifying basis rather than using a predefined basis. This will be evaluated as it has shown promise in other fields. The effect of combining the proposed with existing approaches such as linear discriminant analysis (LDA) will also be examined.

ACKNOWLEDGMENT

Portions of the research in this paper use the CASIA Gait Database collected by Institute of Automation, Chinese Academy of Sciences.

REFERENCES

[1] M. S. Nixon, J. N. Carter, and C. Yam, "Automated person recognition by walking and running via model-based approaches," *IEEE Trans. on Pattern Recognition*, vol. 37, no. 5, pp. 1057–1072, 2004.

[2] J. Han and B. Bhanu, "Individual recognition using gait energy image," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 28, no. 2, pp. 316–322, 2006.

[3] E. Candès and M. Wakin, "An introduction to compressive sampling," *Signal Processing Magazine, IEEE*, vol. 25, no. 2, pp. 21–30, 2008.

[4] D. L. Donoho, "Compressed sensing," *IEEE Trans. on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[5] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," in *IEEE Trans. on Pattern Analysis and Machine Intelligence (PAMI)*, 2008.

[6] P. Nagesh and B. Li, "A compressive sensing approach for expression-invariant face recognition," vol. 0, pp. 1518–1525, 2009.

[7] S. Cotter, "Recognition of occluded facial expressions using a fusion of localized sparse representation classifiers," in *Proc. IEEE Int. Workshop on Digital Signal Processing and Signal Processing Education (DSP/SPE)*, jan. 2011, pp. 437 –442.

[8] Z.-L. Ying, Z.-W. Wang, and M.-W. Huang, "Facial expression recognition based on fusion of sparse representation," in *Proc. IEEE Int. Conf. on Advanced intelligent computing theories and applications*, ser. ICIC'10. Berlin, Heidelberg: Springer-Verlag, 2010, pp. 457–464.

[9] O. Concha, R. Xu, and M. Piccardi, "Compressive sensing of time series for human action recognition," in *Proc. IEEE Int. Conf. on Digital Image Computing: Techniques and Applications (DICTA)*, dec. 2010, pp. 454 –461.

[10] C. Lin and K. Wang, "A behavior classification based on enhanced gait energy image," in *Proc. IEEE Int. Conf. on Networking and Digital Society (ICNDS)*, vol. 2, 2010, pp. 589–592.

[11] X. Huang and N. V. Boulgouris, "Gait recognition using linear discriminant analysis with artificial walking conditions," in *Proc. IEEE Int. Conf. on Image Processing (ICIP)*, 2010, pp. 2461–2464.

[12] E. Candes, "The restricted isometry property and its implications for compressed sensing," *Comptes Rendus Mathematique*, vol. 346, no. 9-10, pp. 589–592, May 2008.

[13] E. Candés, "Compressive sensing," in *Int. Conf. on Congress of Mathematics*, 2006.

[14] D. Chen, S. Denman, C. Fookes, and S. Sridharan, "Accurate silhouettes for surveillance - improved motion segmentation using graph cuts," in *Proc. IEEE Int. Conf. on Digital Image Computing: Techniques and Applications (DICTA)*, 2010, pp. 369–374.

[15] CASIA, "Casia gait database, http://www.sinobiometrics.com," 2009.