

Estudo da Influência da Redução de Dimensionalidade em Recuperação de Conteúdo: Aplicação em Dados Temporais de Exames de Manometria Ano-retal

Everton Alvares Cherman¹, Huei Diana Lee^{1,2}, André Gustavo Maletzke^{1,2},
Carlos Andres Ferrero^{1,2}, João José Fagundes³,
Cláudio Sadi Rodrigues Coy³, Feng Chung Wu^{1,2,3}

¹Centro de Engenharias e Ciências Exatas – Universidade Estadual do Oeste do Paraná
Laboratório de Bioinformática – LABI
Parque Tecnológico Itaipu – PTI
Caixa Postal 39, 85856-970 – Foz do Iguaçu, PR, Brasil

²Instituto de Ciências Matemáticas e de Computação – Universidade de São Paulo
Laboratório de Inteligência Computacional – LABIC
Caixa Postal 668, 13560-970 – São Carlos, SP, Brasil

³Faculdade de Ciências Médicas – Universidade Estadual de Campinas
Serviço de Coloproctologia
Caixa Postal 6111, 13083-970 – Campinas, SP, Brasil

{evertoncherman,hueidianalee}@gmail.com

Abstract. *A great data volume from examinations is registered in the medical area. Frequently, these data contains a time characteristic, as in the ano-rectal manometry examination. Accordingly, time series analysis techniques may be applied intending to help specialists, for example, in the process of decision making. In this work, a content retrieving technique is applied to ano-rectal manometry examination data and the data dimensionality reduction influence is evaluated. There were no statistical significant difference between using original data and reduced dimension data, thus it is possible to summarize data without losing information quality.*

Resumo. *Na área médica é registrado um grande volume de dados referentes a exames de pacientes. Esses dados freqüentemente contêm a característica temporal, como é o caso do exame de manometria ano-retal. Nesse sentido, técnicas computacionais de análise de séries temporais podem ser aplicadas para auxiliar especialistas, por exemplo, no processo de tomada de decisão. Neste trabalho é aplicada a técnica de recuperação de conteúdo sobre os dados de exames de manometria ano-retal e é avaliada a influência da redução de dimensionalidade dos dados. Não foi evidenciada diferença estatisticamente significativa entre a utilização dos dados originais e dos dados com dimensão reduzida, o que possibilita resumir os dados sem perda na qualidade das informações.*

1. Introdução

O avanço tecnológico observado nas últimas décadas tem permitido um grande acúmulo de dados em diversas áreas do conhecimento devido à redução dos custos e do aumento

da capacidade dos meios de armazenamento digitais. A análise detalhada desse grande volume de dados pode trazer informações úteis, como o auxílio para especialistas no processo de tomada de decisão. O apoio computacional para a análise desse grande volume de dados é de grande importância, tendo em vista que a análise manual geralmente é muito custosa ou até inviável [Lee 2005, Lee and Monard 2003]. Processos computacionais, como o de Descoberta de Conhecimento em Bases de Dados – DCBD –, podem auxiliar nessa tarefa [Fayyad et al. 1996].

O DCBD é composto por três etapas: pré-processamento, Mineração de Dados – MD – e pós-processamento. Na primeira etapa é realizada a limpeza, a seleção e a preparação dos dados para um formato adequado à MD. Nessa segunda etapa são aplicadas técnicas de aprendizado de máquina sobre os dados com intuito de identificar padrões e construir modelos representativos aos dados. Na última etapa os modelos construídos são avaliados e validados por especialistas.

Em muitas áreas, grandes volumes de dados são registrados diariamente, e especificamente na área médica são registrados dados a respeito do estado de saúde de pacientes, por exemplo, dados referentes a exames como o eletrocardiograma e a Manometria Anorectal – MA. A MA tem como objetivo diagnosticar pacientes quanto à incontinência fecal, a qual é caracterizada pela perda da capacidade do paciente em controlar a passagem de fezes e gases em tempo e lugar adequados e socialmente aceitáveis. Essa condição apresenta incidência de 2% na população em geral, sendo 66% desses casos em mulheres, e em 17% das pessoas com idade superior aos 70 anos [Saad 2002]. Os dados obtidos nesse exame são considerados dados temporais, isto é, dados que são capturados ao longo do tempo e podem ser definidos como Séries Temporais – ST.

Nesse contexto, técnicas de Análise de Séries Temporais – AST – têm sido propostas na literatura e englobam desde técnicas de pré-processamento dos dados até aplicações como classificação, previsão, *clustering* e recuperação de conteúdo [Morettin and Toloí 2006]. O pré-processamento tem grande influência na qualidade dos dados, o que torna essa fase uma importante tarefa para o sucesso das aplicações em AST. As técnicas de pré-processamento de ST podem ser representadas por métodos de remoção de ruídos, remoção de tendência, remoção de sazonalidade, redução de dimensionalidade e discretização.

Neste trabalho, em andamento, está sendo pesquisada a influência das técnicas de pré-processamento em aplicações de ST, mais especificamente na recuperação de conteúdo de exames de MA. As técnicas de pré-processamento aplicadas neste trabalho referem-se à redução de dimensionalidade, as quais apresentam vantagens como a redução do tempo computacional necessário no processamento dos dados e a ênfase em características de interesse, por exemplo, o comportamento global de uma ST [Chakrabarti et al. 2002].

Este trabalho está organizado da seguinte maneira: na Seção 2 é apresentado o método de captura dos dados do exame de MA e os métodos de redução de dimensionalidade e recuperação de conteúdo utilizados; na Seção 3, os resultados da aplicação da metodologia são apresentados e discutidos; as conclusões e os trabalhos futuros são apresentados na Seção 4.

2. Materiais e Métodos

A manometria ano-retal é um exame imprescindível para o diagnóstico da enfermidade de incontinência fecal. O exame consiste de um cateter axial, o qual é posicionado no esfíncter anal externo do paciente com o intuito de medir a pressão anal de contração. O especialista solicita ao paciente realizar contração voluntária por um período de tempo definido e, com isso, o instrumento realiza a captura de dados. O cateter é constituído de oito canais, os quais são dispostos radialmente, que possibilitam capturar dados de pressão do esfíncter anal externo. Estudos têm sido realizados com o objetivo de extrair, desses dados, características que representem as ST obtidas dos sensores. Essas características, as quais têm como objetivo diferenciar pacientes normais de pacientes com algum grau de incontinência, incluem a pressão máxima de contração voluntária, a capacidade de sustentação, entre outras [Saad 2002]. Neste trabalho, a seqüência temporal dos dados é considerada para análise, o que permite analisar de maneira diferenciada os períodos de construção voluntária do paciente.

Os dados utilizados neste trabalho correspondem a 20 exames MA realizados pelo Serviço de Coloproctologia da UNICAMP no período de Maio/1995 a Novembro/1996. Desse modo, foi construída uma BD, onde cada exame contém as ST obtidas pelos sensores e o diagnóstico realizado pelo especialista. Esse diagnóstico representa a classe de cada exemplo, a qual foi definida em conjunto com especialistas como *Normal*, isto é, o paciente não apresenta incontinência fecal, ou *Grau III*, o paciente apresenta quadro de incontinência fecal de grau III [Saad 2002]. A BD utilizada é constituída de 12 exemplos da classe *Normal* e 8 exemplos da classe *Grau III*. Na Figura 1 é apresentado um exemplo de ST do exame de MA obtido de um dos oito sensores dispostos no canal ano-retal.

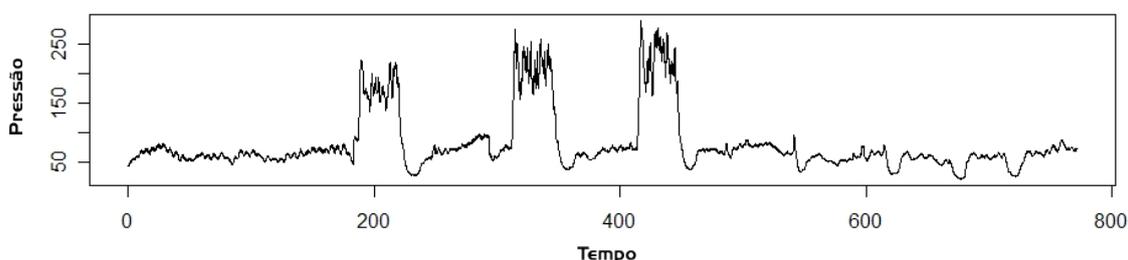


Figura 1. Exemplo de ST do exame de MA obtida de um dos oito sensores dispostos no canal ano-retal.

Para cada exame de MA foi considerado o primeiro período de contração voluntária realizado pelo paciente, o qual tem duração de 40 segundos. A frequência de amostragem dos dados é caracterizada por oito amostras por segundo, e com isso, cada exame de MA é representado por 320 amostras. Neste trabalho, a análise foi realizada utilizando a resultante da pressão exercida pelos oito sensores, a qual foi calculada como a soma dos valores dos oito sensores em cada instante de coleta. É importante salientar que esta abordagem foi utilizada no intuito de simplificar a representação dos dados, devido ao foco deste trabalho direcionar para a avaliação da influência da redução de dimensionalidade na recuperação de conteúdo.

A seguir é apresentada a metodologia proposta neste trabalho para a realização de experimentos, a qual considera séries temporais rotuladas e está constituída de três etapas:

(1) redução de dimensionalidade, (2) recuperação de conteúdo, (3) análise de resultados. Essas etapas são apresentadas a seguir.

2.1. Redução de Dimensionalidade

Dentre as análises realizadas em séries temporais, pesquisas relacionadas à busca por similaridade (recuperação de conteúdo) em bases de dados de ST, podem fornecer uma ferramenta útil para a análise exploratória de dados e também um importante elemento para diversas aplicações de mineração de séries temporais, tais como *clustering*, classificação e mineração por associação de regras [Chakrabarti et al. 2002]. O cálculo da similaridade entre duas séries consiste na utilização de uma medida, como a distância Euclidiana, com a finalidade de quantificar a diferença existente entre essas séries. No entanto, bases de dados com centenas de gigabytes são comuns e um grande volume de dados submetido a esse cálculo pode trazer problemas relacionados ao tempo necessário para a realização desse processo. Nesse contexto, os métodos mais promissores de busca por similaridade estão relacionados a técnicas que oferecem redução da dimensionalidade, isto é, técnicas que reduzem o espaço necessário para se armazenar ST preservando ao máximo as características originais dos dados. De acordo com esses fatores, o objetivo é representar uma série temporal de dimensão n , em uma dimensão w , onde $w \ll n$, após a aplicação dos métodos de redução de dimensionalidade.

Piecewise Aggregate Approximation

O método *Piecewise Aggregate Approximation* – PAA – constitui um dos métodos de redução de dimensionalidade mais utilizados. Esse método consiste na separação da ST em pequenos segmentos uniformes quanto ao tamanho, e para cada segmento é atribuído um valor que irá representá-lo [Lin et al. 2003]. Esse valor corresponde à média dos valores contidos nesse segmento. Formalmente, uma série temporal C de tamanho n pode ser representada em um espaço de w dimensões por um vetor $\bar{C} = \bar{c}_1, \dots, \bar{c}_w$. O i -ésimo elemento de \bar{C} é calculado pela Equação 1.

$$\bar{c}_i = \frac{w}{n} \sum_{j=\frac{n}{w}(i-1)+1}^{\frac{n}{w}i} c_j \quad (1)$$

Na Figura 2 é apresentado um exemplo da utilização do método PAA, no qual a ST C é representada por oito segmentos.

2.2. Recuperação de Conteúdo

Uma das tarefas de interesse em análise de séries temporais é a recuperação de conteúdo que consiste em recuperar da base de dados os registros mais similares a um exemplar fornecido como consulta. Porém, no contexto de séries temporais, essa tarefa é custosa devido à dimensionalidade dos dados. Por esse motivo, é desejável a realização da redução de dimensionalidade com o objetivo de tornar menos custosa a tarefa de recuperação de conteúdo.

A recuperação de conteúdo tem como objetivo identificar as k ST, contidas em uma base de dados, mais semelhantes à série Q fornecida e recuperá-las. Para isso é

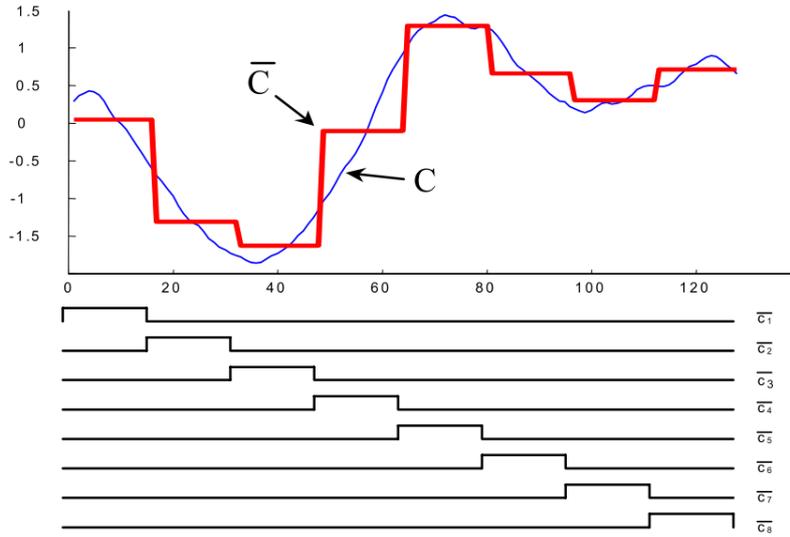


Figura 2. Representação de uma série temporal de tamanho 128 reduzida para 8 dimensões [Lin et al. 2003]

utilizado o método *k-Nearest-Neighbor* – kNN – que consiste em identificar os k elementos mais próximos, considerando uma medida de distância, dado um elemento de entrada [Mörchen 2006]. Um aspecto importante é a definição da medida de similaridade entre ST. Diversas medidas podem ser aplicadas, as quais variam de acordo com as características da série e com o modo como essa série é representada [Last et al. 2004].

A distância Euclidiana é a medida mais utilizada para a comparação de séries temporais [Last et al. 2004, Keogh and Kasetty 2003]. Essa medida permite determinar o comprimento da linha reta entre dois pontos, os quais pertencem a um espaço de m dimensões, onde m corresponde ao tamanho de uma determinada ST. A distância Euclidiana entre duas ST $U(u_1, u_2, \dots, u_m)$ e $F(f_1, f_2, \dots, f_m)$, é definida conforme a Equação 2.

$$D_{UF} = \sqrt{(u_1 - f_1)^2 + (u_2 - f_2)^2 + \dots + (u_m - f_m)^2} = \sqrt{\sum_{i=1}^m (u_i - f_i)^2} \quad (2)$$

onde i indica cada dimensão em cada iteração para o cálculo da distância. Assim, essa medida de similaridade é determinada por meio do cálculo da raiz quadrada da soma dos quadrados da diferença entre os valores da ST.

Embora a distância Euclidiana seja uma das medidas de similaridade mais utilizadas, essa medida apresenta algumas desvantagens quando aplicada ao contexto de ST. Dentre as desvantagens identificadas, a principal delas refere-se à alta sensibilidade na comparação de ST que possuem defasagem uma em relação a outra. Desse modo, com o intuito de amenizar essa deficiência apresentada pela utilização da distância Euclidiana foi proposta a medida denominada *Dynamic Time Warping* – DTW [Chu et al. 2002].

Para se determinar a DTW, dada duas séries temporais U e F de tamanho n_u e n_f , respectivamente, deve-se alinhar ambas as séries. Para isso é construída uma matriz $n_u \times n_f$, na qual os elementos (i, j) da matriz contém a distância $D(u_i, f_j)$ entre dois pontos u_i e f_j e freqüentemente é utilizada a distância Euclidiana para se determinar a distância

entre ambos os pontos. Cada elemento (i, j) da matriz corresponde ao alinhamento entre os pontos u_i e f_j . Essa situação é ilustrada na Figura 3, na qual a rota R é um conjunto contínuo de elementos da matriz que define o mapeamento entre U e F . Devido a essas características, o DTW pode apresentar para alguns problemas um desempenho superior à distância Euclidiana, porém possui algumas desvantagens causadas pela ordem quadrática da matriz de distâncias, como o tempo $O(n^2)$, onde n é o tamanho da ST, e o espaço em memória necessário [Chu et al. 2002].

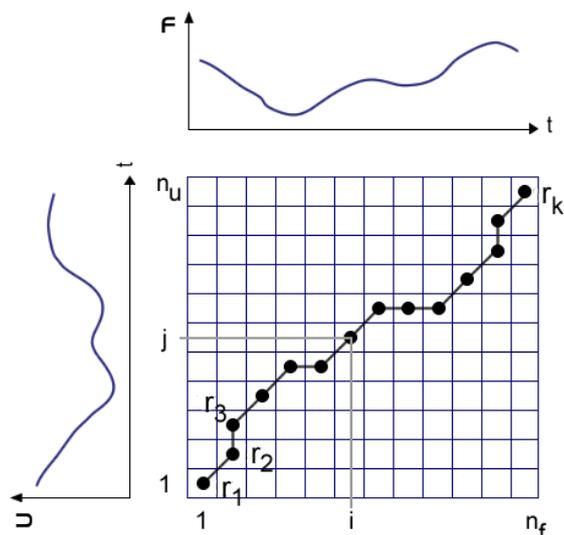


Figura 3. Exemplo de uma matriz obtida a partir da aplicação da DTW [Chu et al. 2002].

2.3. Análise de Resultados

A aplicação da recuperação de conteúdo deve ser avaliada para evidenciar a qualidade das recuperações realizadas, no intuito de comparar diferentes abordagens de recuperação. Uma das técnicas de análise de resultados consiste na construção do gráfico *Precision* \times *Recall*. Esse gráfico tem como objetivo avaliar a relevância do conteúdo recuperado. Para isso, cada exemplar da BD é retirado e fornecido como argumento de consulta e um valor de k é determinado para o método kNN. Desse modo, são recuperados os k exemplares mais próximos ao exemplar fornecido como consulta e em seguida, é possível calcular a *precision* e o *recall* dos exemplares recuperados.

Os valores de *precision* e de *recall* de uma consulta para um determinado k são calculados por meio das Equações 3 e 4, respectivamente:

$$Precision = \frac{ERR}{TER} \quad (3)$$

$$Recall = \frac{ERR}{TE} \quad (4)$$

onde, ERR corresponde ao número de exemplos relevantes recuperados, isto é, da mesma classe do exemplar fornecido como consulta; TER , ao total de exemplos recuperados; e,

TE , ao total de exemplos contidos na base de dados que contém a mesma classe do exemplar fornecido com consulta.

No entanto, para se obter uma aproximação aos valores verdadeiros de *precision* e de *recall* da recuperação de conteúdo, não deve ser realizada a recuperação utilizando apenas um exemplar como argumento de consulta, mas uma quantidade de exemplares suficientemente representativa. Para isso, uma técnica que pode ser utilizada é a de *Leave-One-Out Cross-validation* [Rezende 2003].

3. Resultados e Discussão

A metodologia apresentada foi aplicada em 20 exames de MA. De acordo com a metodologia, na Etapa 1, foi realizada a redução de dimensionalidade dos dados por meio da utilização do PAA com um fator de redução de 90%. Desse modo, foram considerados dois conjuntos de séries temporais: o conjunto de ST originais, contendo 100% dos pontos, e o conjunto das ST resultantes da redução de dimensionalidade, contendo 10% dos pontos. Esses dois conjuntos de dados foram submetidos à Etapa 2, de recuperação de conteúdo, utilizando as medidas de distância Euclidiana e DTW. Na Tabela 1, são apresentadas as combinações dos conjuntos experimentais submetidos à tarefa de recuperação de conteúdo.

Id	Medida de Similaridade	Dimensão
1	Euclidiana	100%
2	Euclidiana	10%
3	DTW	100%
4	DTW	10%

Tabela 1. Combinações de medidas de similaridade aplicadas e a porcentagem de pontos utilizada em relação às ST originais.

Na Etapa 3, foi realizada a avaliação dos resultados utilizando a técnica de *Precision* \times *Recall*. Na Tabela 2 são apresentados os dados de *precision* para cada valor de *recall* observados na utilização das diferentes combinações de medidas de similaridade e dimensões dos dados, bem como os valores de média e desvio-padrão. Também nessa tabela, são apresentados os *p-valores* referente à aplicação do teste estatístico *t-student* para a verificação de diferença significativa entre as colunas com e sem redução de dimensionalidade para cada técnica.

Na Figura 4 é apresentado o resultado utilizando essa técnica, para a classe *Normal* e para a classe *Grau III*. A avaliação desse gráfico pode ser realizada por meio da comparação visual entre as diferentes medidas de similaridade e também entre as diferentes dimensões utilizadas. Como o foco desse trabalho é avaliar a influência da redução de dimensionalidade na qualidade da recuperação de conteúdo, não serão realizadas comparações entre as diferentes medidas de similaridade utilizadas.

Nesse sentido, a análise dos gráficos de ambas as classes é realizada separadamente e em cada gráfico são comparados os valores de *precision* da recuperação de conteúdo entre as ST com 100% e 10% dos pontos. Em uma análise visual do gráfico relacionado à classe normal, na utilização da distância Euclidiana como medida de similaridade, não é possível observar diferença entre os dados de diferentes dimensões. Para

Recall (%)	Precision da classe							
	Normal				Grau III			
	1	2	3	4	1	2	3	4
10	66,67	66,67	58,33	58,33	37,50	37,50	37,50	37,50
20	56,48	56,48	61,11	71,53	47,92	47,92	40,63	39,93
30	58,33	58,33	62,50	66,88	43,75	46,88	38,75	37,50
40	55,85	55,85	66,10	60,81	41,04	41,04	36,46	36,46
50	56,89	56,89	59,78	58,33	42,86	41,07	37,50	37,50
60	57,95	57,95	59,51	57,92	45,57	45,57	38,75	38,04
70	59,03	59,03	61,11	61,11	39,88	39,62	35,27	35,58
80	59,94	59,94	62,04	62,02	37,98	37,53	36,67	35,77
90	56,81	56,81	58,18	58,53	35,29	35,29	35,38	34,31
100	56,73	56,73	56,73	56,73	39,76	39,76	36,84	38,34
Média	58,47	58,47	60,54	61,22	41,15	41,22	37,37	37,09
Desvio-padrão	3,15	3,15	2,67	4,64	3,89	4,26	1,65	1,60
p-valor	—		0.6178		0.8731		0.3055	

Tabela 2. Tabela com os valores de *precision* das quatro combinações de cada classe. Cada linha da tabela corresponde a um valor de *recall*.

DTW como medida de similaridade, é possível observar um comportamento semelhante entre as duas dimensões, de modo que não foi observada diferença visual nos valores de *precision* quando os valores de *recall* correspondem a 10% e 70% a 100%.

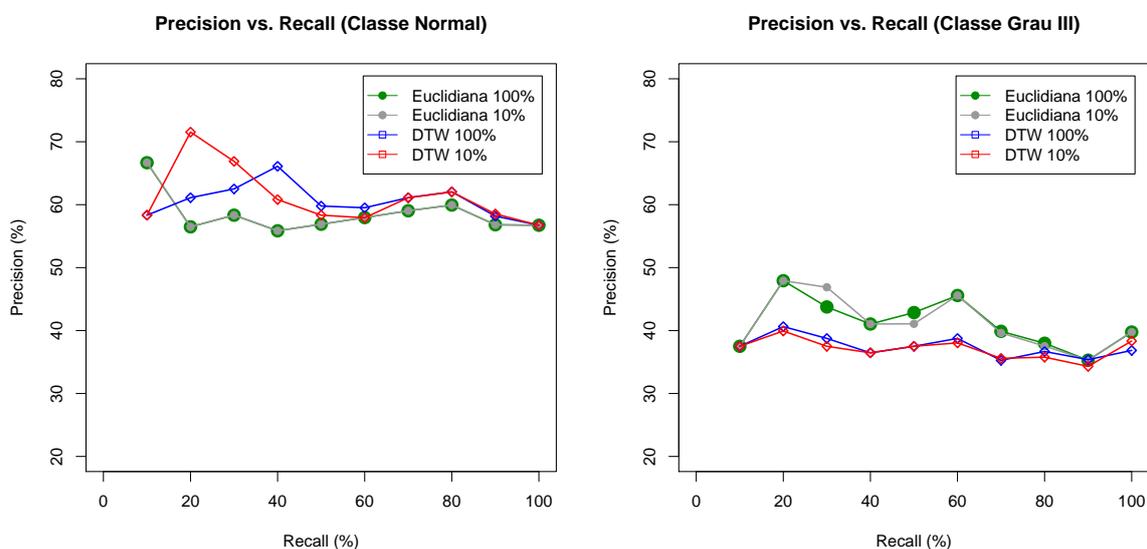


Figura 4. Gráficos *Precision* \times *Recall* para exemplares da classe *Normal* e *Grau III*.

No gráfico *Precision* \times *Recall* de exemplos da classe *Grau III* também pode ser constatada a baixa diferença visual entre as duas dimensões de ST utilizadas. Ao considerar a medida Euclidiana, somente nos valores de precisão para 30% e 50% de *recall* é observada diferença. Para o DTW como medida de similaridade, ambas as curvas de *Precision* \times *Recall* apresentam comportamentos visualmente similares, com pequenas variações em alguns pontos.

Como apresentado na Tabela 2, foi realizada uma análise estatística para verificar a diferença do desempenho de recuperação de conteúdo utilizando, ou não, redução de dimensionalidade. Nesse sentido, foi aplicado o teste estatístico *t-student* para dados emparelhados com nível de significância de 95% para verificar a existência de diferença estatística entre os quatro pares de dados em análise. Por meio da aplicação do teste foi possível evidenciar que não existe diferença estatística entre todas as combinações para um nível de significância de 95%. Esse resultado é considerado positivo, pois através da utilização de dados com apenas 10% da dimensão original é possível prover uma qualidade na recuperação de conteúdo semelhante à utilização de 100% dos dados. Essa característica é importante principalmente para técnicas computacionais que contém complexidade computacional de tempo e espaço elevada, como é o caso do DTW, o qual contém complexidade quadrática em ambos os fatores.

4. Conclusões

Neste trabalho, em andamento, foi apresentado um estudo inicial sobre a influência da redução de dimensionalidade na aplicação da técnica de recuperação de conteúdo em dados temporais de MA. A redução de dimensionalidade foi obtida por meio do método PAA, que é considerado um dos métodos mais simples e freqüentemente utilizado, e conforme observado nos resultados, o PAA possibilitou a redução da dimensão das ST sem que houvesse perda significativa da precisão na aplicação da técnica de recuperação de conteúdo. Essa característica também foi observada estatisticamente, de modo que não foi possível verificar diferença estatística entre as precisões das ST de dimensões originais e reduzidas. Desse modo, foi possível armazenar as ST contidas nas BD, com apenas 10% do espaço total necessário para armazenar as ST originais, e também, como consequência da redução dos pontos que representam as ST, é possível realizar o processamento das ST com um menor custo de tempo, o que pode ser considerado positivo, principalmente com a aplicação de DTW como distância de similaridade, a qual contém complexidade de tempo de ordem quadrática.

Portanto, a utilização da técnica PAA para redução de dimensionalidade sobre os dados de MA considerados neste trabalho influenciaram positivamente na aplicação de recuperação de conteúdo do ponto de vista da redução do tempo gasto no processamento e do espaço necessário para o armazenamento das ST.

Como trabalhos futuros pretende-se avaliar a redução de dimensionalidade considerando outras técnicas e outros fatores de redução e também a utilização de técnicas de discretização sobre os dados. Outros trabalhos futuros consistem na aplicação de outras abordagens para o pré-processamento dos dados dos exames de MA, como a utilização das ST dos oito canais separadamente e a consideração de todos os momentos de contração voluntária do exame.

Agradecimentos

Ao Programa de Desenvolvimento Tecnológico Avançado do Parque Tecnológico Itaipu (PDTA/FPTI-BR) pelo apoio financeiro.

Referências

Chakrabarti, K., Keogh, E., Mehrotra, S., and Pazzani, M. (2002). Locally adaptive dimensionality reduction for indexing large time series databases. *ACM Trans. Database*

Syst., 27(2):188–228.

- Chu, S., Keogh, E. J., Hart, D., and Pazzani, M. J. (2002). Iterative deepening dynamic time warping for time series. In *Second SIAM International Conference on Data Mining*, pages 1–18, Arlington, Virginia, USA.
- Fayyad, U., Piatetsky, G., and Smyth, P. (1996). From data mining to knowledge discovery in databases. *AI Magazine*, 17:37–54.
- Keogh, E. and Kasetty, S. (2003). On the need for time series data mining benchmarks: A survey and empirical demonstration. volume 7, pages 349–371, Hingham, MA – USA. Kluwer Academic Publishers.
- Last, M., Kandel, A., and Bunke, H. (2004). *Machine perception and artificial intelligence: data mining in time series databases*. World Scientific, Massachusetts – USA.
- Lee, H. D. (2005). *Seleção de atributos importantes para a extração de conhecimento de bases de dados*. PhD thesis, Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – Brasil.
- Lee, H. D. and Monard, M. C. (2003). Seleção de atributos para algoritmos de aprendizado de máquina supervisionado utilizando como filtro a dimensão fractal. *Revista de La Sociedad Chilena de Ciencia de La Computación*, pages 1–8.
- Lin, J., Keogh, E., Lonardi, S., and Chiu, B. (2003). A symbolic representation of time series, with implications for streaming algorithms. In *DMKD '03: Proceedings of the 8th ACM SIGMOD workshop on Research issues in data mining and knowledge discovery*, pages 2–11, New York, NY – USA. ACM.
- Mörchen, F. (2006). Time series knowledge mining. Master's thesis, Philipps-Universität Marburg, Marburg – Germany.
- Morettin, P. A. and Toloí, C. M. C. (2006). *Análise de séries temporais*. Edgard Blücher, São Paulo – Brasil.
- Rezende, S. O. (2003). *Sistemas Inteligentes: Fundamentos e Aplicações*. Editora Manole, Barueri – Brasil.
- Saad, L. H. C. (2002). *Quantificação da função esfinteriana pela medida da capacidade de sustentação da pressão de contração voluntária do canal anal*. PhD thesis, Universidade Estadual de Campinas.