



República Federativa do Brasil  
Ministério da Economia  
Instituto Nacional da Propriedade Industrial

(21) BR 102018006344-8 A2



(22) Data do Depósito: 28/03/2018

(43) Data da Publicação Nacional: 15/10/2019

(54) **Título:** APROXIMAÇÃO PARA A TRANSFORMADA DE KARHUNEN-LOÈVE COM BAIXO CUSTO ARITMÉTICO COM APLICAÇÕES EM COMPRESSÃO DE IMAGENS

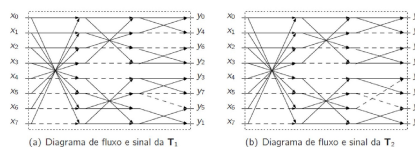
(51) **Int. Cl.:** H04N 19/42; H04N 19/60.

(52) **CPC:** H04N 19/42; H04N 19/60.

(71) **Depositante(es):** UNIVERSIDADE FEDERAL DE PERNAMBUCO; UNIVERSIDADE FEDERAL DE SANTA MARIA.

(72) **Inventor(es):** ANABETH PETRY RANDÜZ; RENATO JOSÉ DE SOBRAL CINTRA; FÁBIO MARIANO BAYER; DIEGO RAMOS CANTERLE.

(57) **Resumo:** A transformada de Karhunen-Loève (KLT) é frequentemente utilizada na descorrelação de dados e na redução de dimensionalidade. Por sua característica de compactação de energia em poucos componentes, encontram-se aplicações importantes da KLT no contexto de compressão de imagens. Contudo, seu uso em aplicações reais é impossibilitado devido ao alto custo computacional requerido para seu cálculo a cada sinal de entrada. Neste contexto, a presente invenção propõe uma versão de baixo custo computacional para a KLT por meio da aplicação da função sinal nos elementos da matriz induzida pela KLT. Denominamos a transformada proposta por KLT sinalizada (SKLT). É dada ênfase nas transformadas de comprimento 8 devido à utilização desse comprimento de transformadas na codificação de imagem, como no padrão JPEG. A SKLT foi numericamente avaliada por meio de medidas de codificação e de distância totais para a KLT exata.



## **APROXIMAÇÃO PARA A TRANSFORMADA DE KARHUNEN-LOÈVE COM BAIXO CUSTO ARITMÉTICO COM APLICAÇÕES EM COMPRESSÃO DE IMAGENS**

1. Propõem-se uma nova transformada que aproxima a transformada de Karhunen-Loève de dimensão  $8 \times 8$  e aplica-se a transformada em blocos. Esse tipo de transformada pode ser utilizada em compressão de imagens e vídeo como nos padrões JPEG e HEVC. Para aplicação substitui-se a KLT exata pela aproximação proposta quando o sinal segue um processo markoviano de primeira ordem, tendo um alto poder de decorrelação dos dados. Imagens estacionárias são um exemplo deste tipo de sinal. A transformada proposta possui um baixo custo computacional, o que potencializa sua aplicação em compressão de imagens.

2. A transformada de Karhunen-Loève (KLT) é comumente utilizada no contexto de decorrelação de dados e redução de dimensionalidade. Ela envolve um procedimento matemático que transforma as variáveis correlacionadas em variáveis não-correlacionadas. Usualmente, apenas as primeiras componentes principais são utilizadas para representar grande parte da variabilidade (energia) dos dados (sinal). Essa característica de retenção de energia tem grande importância para a compressão de imagens, pois a KLT é capaz de preservar as informações da imagem mesmo reduzindo a dimensionalidade dos dados.

3. Quando o sinal de entrada é um processo Markoviano de primeira ordem, a KLT decorrelaciona completamente o sinal no domínio da transformada. De fato, a KLT é uma transformada linear ótima capaz de minimizar o erro quadrático médio na compressão dos dados e concentrar maior energia em poucos coeficientes do sinal de saída. No entanto, o uso da KLT em processamento de sinais é reduzido, pois sua construção

depende da matriz de variância e covariância do sinal de entrada, dificultando o desenvolvimento de algoritmos rápidos que a implemente. Neste contexto, surge a transformada discreta do cosseno (DCT) que é uma aproximação assintótica para a KLT quando a correlação do sinal de entrada tende à unidade. A definição da DCT independe do sinal de entrada, o que permite o desenvolvimento de algoritmos rápidos computacionalmente eficientes. Assim, a DCT é amplamente adotada em padrões de compressão de imagem e vídeo como JPEG, MPEG e HEVC, por exemplo.

4. Entretanto, o uso da DCT ainda pode ser proibitivo em contextos que disponham de baixo poder de processamento ou severa restrição de autonomia energética. Isto se dá pelo fato de que a DCT, em sua forma exata, requer multiplicações por números reais que exigem aritmética de ponto flutuante para sua implementação. Implementações em hardware para realizar operações aritméticas em ponto flutuante—especialmente multiplicações—são usualmente demandantes em termos de circuitaria e de consumo. Neste sentido, diversas aproximações para a DCT vêm sendo propostas, de forma que sejam livres de multiplicação. Uma importante aproximação para a DCT é a clássica DCT sinalizada (SDCT). A SDCT é construída através da aplicação da função sinal nos elementos da matriz da DCT, se tornando livre de multiplicações não triviais. Dessa forma, a SDCT possui capacidade de codificação semelhante à DCT, mas oferece um custo computacional bastante reduzido.

5. Considerando as boas propriedades da KLT e – ao nosso melhor conhecimento – ausência de algoritmos e dispositivos no que tange alternativas de baixo custo para a KLT, objetivamos construir aproximações para a KLT seguindo uma variação da abordagem utilizada para derivar a SDCT, que emprega a função sinal como meio para aproximação matricial.

6. A KLT é uma transformação linear representada por uma matriz orthogonal  $K_N^{(\rho)}$  que descorrelaciona um sinal N-dimensional de entrada  $\mathbf{x} = [x_0 \ x_1 \ \dots \ x_{N-1}]^T$  resultando em um sinal descorrelacionado  $\mathbf{y} = [y_0 \ y_1 \ \dots \ y_{N-1}]^T$ . Considerando que  $\mathbf{x}$  seja um processo Markoviano de primeira ordem com autocorrelação  $\rho$ , assumido na maioria das aplicações em processamento de sinais, esta operação é realizada matricialmente por

$$\mathbf{y} = \mathbf{K}_N^{(\rho)} \cdot \mathbf{x}.$$

7. Considerando o sinal  $\mathbf{x}$  com a matriz de covariância associada dada por:

$$\mathbf{\Sigma}_N = E\{(\mathbf{x} - E\{\mathbf{x}\})(\mathbf{x} - E\{\mathbf{x}\})^T\},$$

temos que as linhas correspondentes da matriz orthogonal da transformada de Karhunen-Loève  $K_N^{(\rho)}$  são os autovetores da matriz de covariância simétrica positiva  $\mathbf{\Sigma}_N$ . Uma propriedade essencial dessa transformada é que ela produz coeficientes descorrelacionados, ou seja:

$$E\{(\mathbf{y} - E\{\mathbf{y}\})(\mathbf{y} - E\{\mathbf{y}\})^T\} = \mathbf{K}_N^{(\rho)} \cdot \mathbf{\Sigma}_N \cdot \mathbf{K}_N^{(\rho)T} = \begin{bmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_N \end{bmatrix},$$

em que cada  $\lambda_i$  representa um autovalor da matriz  $\mathbf{\Sigma}_N$ .

8. A matriz de covariância de  $\mathbf{x}$  é dada pelos seguintes elementos:

$$r_{i,j}^{(x)} = \rho^{|i-j|},$$

em que  $\rho$  é o coeficiente de correlação,  $0 < \rho < 1$ , e  $i, j = 0, 1, \dots, N-1$ . Os  $(i,j)$ -ésimos elementos da matriz de transformação  $K_N^{(\rho)}$ , dado um valor arbitrário de  $\rho$ , são dados por:

$$u_{ij} = \sqrt{\frac{2}{N + \lambda_j}} \sin \left[ \omega_j \left( i - \frac{N-1}{2} \right) + \frac{(j+1)\pi}{2} \right],$$

em que os autovalores da matriz de covariância são:

$$\lambda_j = \frac{1 - \rho^2}{1 + \rho^2 - 2\rho \cos \omega_j},$$

e  $\omega_1, \omega_2, \dots, \omega_N$  são as  $N$  soluções da equação:

$$\tan N\omega = \frac{-(1 - \rho^2) \sin \omega}{(1 + \rho^2) \cos \omega - 2\rho}.$$

9. No contexto de processamento de imagens, os dados dos pixels adjacentes são altamente correlacionados.

10. A matriz de transformação associada à SKLT é definida por:

$$\mathbf{T}_N^{(\rho)} \triangleq \frac{1}{\sqrt{N}} \text{sign} \left( \mathbf{K}_N^{(\rho)} \right),$$

em que a função  $\text{sign}(\cdot)$  é dada por:

$$\text{sign}(x) = \begin{cases} 1, & \text{se } x > 0, \\ 0, & \text{se } x = 0, \\ -1, & \text{se } x < 0. \end{cases}$$

Quando aplicada a uma matriz, a função sinal é empregada elemento-a-elemento.

11. A proposta desta transformada, assim como a SDCT, se dá principalmente pelo motivo de ela reduzir o número total de operações aritméticas necessárias para seu cálculo. Como é uma transformada binária com entradas  $\pm 1$ , requer apenas adições para sua implementação. Focando na compressão de imagens e na baixa complexidade da SKLT, serão avaliadas apenas transformadas de comprimento  $N = 8$ .

12. Como a SKLT depende de  $N$  e de  $\rho$ , fixamos  $N = 8$  e procedemos uma busca na família de matrizes  $\mathbf{T}_8^\rho$  para diferentes valores de  $\rho$ . Para isso foi feita uma busca computacional variando o  $\rho$  de 0.01 a 0.999 com passos de  $10^{-3}$ . A busca computacional retorna somente duas diferentes matrizes. A primeira,  $\mathbf{T}_1$ , é definida para qualquer valor de autocorrelação no intervalo  $0 < \rho < 0.619$ . A segunda,  $\mathbf{T}_2$ , é definida em  $0.620 < \rho < 1$ . Destaca-se que a  $\mathbf{T}_2$  é a mesma matriz da SDCT. Já a  $\mathbf{T}_1$ , ao nosso melhor conhecimento, é uma transformada nova não encontrada na literatura.

13. As respectivas matrizes da  $\mathbf{T}_1$  e da  $\mathbf{T}_2$  são dadas por:

$$\mathbf{T}_1 = \frac{1}{\sqrt{8}} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & 1 & -1 & -1 & 1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \end{bmatrix}$$

e

$$\mathbf{T}_2 = \frac{1}{\sqrt{8}} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & 1 & -1 & -1 & 1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \end{bmatrix}.$$

14. Além de buscar transformadas livres de multiplicação, é crucial o desenvolvimento de algoritmos rápidos capazes de reduzir o custo aritmético do cômputo das transformadas. Assim, propomos para a  $\mathbf{T}_1$  o algoritmo

rápido que está demonstrado na forma de diagrama de fluxo de sinal (DFS) na Figura 1(a); e para a  $T_2$  na Figura 1(b). Nesses DFS, o encontro de duas setas indica uma adição, desta forma é fácil verificar que as transformadas necessitam 24 adições cada para o seu computo.

15. As transformadas  $T_1$  e  $T_2$  são não ortogonais, mas suas inversas são de baixa complexidade pois possuem como entradas apenas elementos do conjunto  $\{\pm 1, \pm 2\}$  e são dadas, respectivamente, por:

$$T_1^{-1} = 8\sqrt{8} \begin{bmatrix} 1 & 0 & 1 & 2 & 1 & 2 & 1 & 0 \\ 1 & 2 & 1 & -0 & -1 & -2 & -1 & -0 \\ 1 & 2 & -1 & -2 & -1 & -2 & 1 & 2 \\ 1 & -0 & -1 & -0 & 1 & 2 & -1 & -2 \\ 1 & -0 & -1 & -0 & 1 & -2 & -1 & 2 \\ 1 & -2 & -1 & 2 & -1 & 2 & 1 & -2 \\ 1 & -2 & 1 & -0 & -1 & 2 & -1 & -0 \\ 1 & -0 & 1 & -2 & 1 & -2 & 1 & -0 \end{bmatrix}$$

e

$$T_2^{-1} = 8\sqrt{8} \begin{bmatrix} 1 & 2 & 1 & 2 & 1 & 0 & 1 & 0 \\ 1 & 2 & 1 & -0 & -1 & -2 & -1 & -0 \\ 1 & -0 & -1 & -2 & -1 & -0 & 1 & 2 \\ 1 & 0 & -1 & 0 & 1 & 2 & -1 & -2 \\ 1 & -0 & -1 & -0 & 1 & -2 & -1 & 2 \\ 1 & -0 & -1 & 2 & -1 & -0 & 1 & -2 \\ 1 & -2 & 1 & -0 & -1 & 2 & -1 & -0 \\ 1 & -2 & 1 & -2 & 1 & -0 & 1 & -0 \end{bmatrix}$$

16. Para a transposta de  $T_1^{-1}$  temos o algoritmo rápido que está demonstrado na Figura 2(a); e para a transposta de  $T_2^{-1}$  na Figura 2(b). Multiplicações por  $\pm 2$  são *bit-shifts* e são considerados de baixa complexidade computacional. Verifica-se que para o computo de  $T_1^{-1}$  são necessárias 22 adições e 4 *bit-shift* e para o computo de  $T_2^{-1}$  são necessárias 20 adições e 4 *bit-shift*.

17. Avaliamos o desempenho das transformadas propostas quando aplicadas na compressão de imagens. Por simplicidade, mas sem perda de generalidade, serão consideradas imagens de 8 bits em escala de cinza. Uma imagem digital pode ser definida como uma matriz, em que cada pixels representa um nível de escala de cinza com valores de zero à  $255 = 2^8 - 1$ .

18. A compressão de imagens do tipo JPEG segue os seguintes passos :  
 (i) a imagem é dividida em blocos  $A_j$  de ordem  $8 \times 8$ ; (ii) é aplicada a transformação bidimensional  $T$  em cada bloco, por  $B_j = T \cdot A_j \cdot T_1^{-1}$ ; (iii) empregando a sequência zig-zag padrão retém-se apenas os  $r$  coeficientes iniciais em cada bloco  $B_j$  e os  $64 - r$  coeficientes restantes são anulados, obtendo os blocos truncados  $\bar{B}_j$  (nesta etapa acontece a compressão); (iv) aplica-se a transformada bidimensional inversa, dada por  $\bar{A}_j = T^{-1} \cdot \bar{B}_j \cdot T$  e (v) recompõem-se os blocos comprimidos  $\bar{A}_j$  no lugar dos blocos originais da imagem  $A_j$ . A imagem final comprimida  $\bar{A}$  pode então ser comparada com a imagem original  $A$  para avaliar a perda de qualidade imposta pela compressão.

19. Utilizamos a relação sinal-ruído de pico (PSNR) como figura de mérito para esta avaliação. A PSNR é dado por:

$$\text{PSNR} = 10 \cdot \log_{10} \left( \frac{\text{MAX}^2}{\text{EQM}} \right),$$

em que

$$\text{EQM} = \frac{\sum_{i=1}^M \sum_{j=1}^M (A_{ij} - \bar{A}_{ij})^2}{M^2},$$

$A$  é a imagem original,  $\bar{A}$  a imagem reconstituída,  $M$  é a dimensão de ambas imagens,  $\text{MAX} = 255$  é o valor máximo dos elementos de  $A$ .

20. A Figura 4 apresenta os valores de PSNR para diferentes escolhas de coeficiente retidos  $r$  ( $1 \leq r \leq 45$ ) na compressão da padrão *Lena*, considerando a DCT, a KLT ( $\rho = 0.95$ ),  $T_1$  e  $T_2$ . Como esperado, por ter menor complexidade, as aproximadas propostas perdem um pouco de desempenho, apresentando valores de PSNR um pouco abaixo que a KLT e

DCT. Contudo, ainda dentro de valores de PSNR admissíveis em várias aplicações em compressão de imagens.

21. Como uma avaliação qualitativa, a Figura 3 mostra o resultado da compressão de imagens aplicada na imagem *Lena* retendo apenas os  $r = 10$  primeiros coeficientes espectrais. A Figura 3(a) apresenta a imagem original e as Figuras 3(b), 3(c), 3(d) e 3(e) mostram, respectivamente, as imagens comprimidas utilizando a DCT, a KLT, a  $T_1$  e a  $T_2$ . Qualitativamente, não há grande diferença entre as imagens comprimidas.

22. A Tabela 1 apresenta as complexidades aritméticas das transformadas de comprimento oito com e sem algoritmo rápido. São apresentadas as quantidades de multiplicações e adições requeridas por cada algoritmo. Pode-se observar que as transformadas que não são aproximações exigem multiplicações, já as transformadas aproximadas têm complexidade multiplicativa nula. Observa-se também que as transformadas quando calculadas pelos seus algoritmos rápidos possuem suas complexidades aritméticas reduzidas. Cabe ressaltar que como a KLT depende do sinal de entrada, se tornando diferente para cada entrada, é inviável a construção de algoritmos rápidos para ela.

| Transformada               | Mult. | Ad. |
|----------------------------|-------|-----|
| KLT                        | 64    | 56  |
| DCT sem algoritmo rápido   | 64    | 56  |
| DCT com algoritmo rápido   | 11    | 29  |
| $T_1$ sem algoritmo rápido | 0     | 56  |
| $T_1$ com algoritmo rápido | 0     | 24  |
| $T_2$ sem algoritmo rápido | 0     | 56  |
| $T_2$ com algoritmo rápido | 0     | 24  |

Tabela 1

23. Neste documento foi proposta uma abordagem de baixa complexidade computacional para descorrelação de dados. Foi considerada a função sinal dos elementos da matriz da KLT, ou KLT, para derivar transformadas inteiras de baixo custo aritmético. A complexidade computacional desta nova transformada é consideravelmente baixa, pois independe do sinal de entrada e é livre de multiplicações. A SKLT proposta mostrou-se uma boa aproximação para KLT nas medidas de qualidade elegidas. Ainda, no contexto de compressão de imagens a SKLT se mostra bastante útil, evidenciando imagens adequadas a um custo computacional associado bastante baixo.

## **REIVINDICAÇÕES**

1. APROXIMAÇÃO PARA A TRANSFORMADA DE KARHUNEN-LOÈVE COM BAIXO CUSTO ARITMÉTICO caracterizada pela compressão e descompressão de imagens e vídeo que seja executado sem operações em ponto flutuante utilizando a SKLT no bloco de transformação.
2. APROXIMAÇÃO PARA A TRANSFORMADA DE KARHUNEN-LOÈVE COM BAIXO CUSTO ARITMÉTICO caracterizada por utilizar a transformada inversa proposta ( $SKLT^{-1}$ ) no bloco de transformação 2-D inversa no esquema geral de descompressão de imagem.
3. APROXIMAÇÃO PARA A TRANSFORMADA DE KARHUNEN-LOÈVE COM BAIXO CUSTO ARITMÉTICO caracterizada por executar operações em ponto flutuante e substituir qualquer circuito esquemático rápido para a KLT pelo circuito esquemático rápido para a transformada direta proposta (SKLT) ilustrado na Figura 1.

## FIGURAS

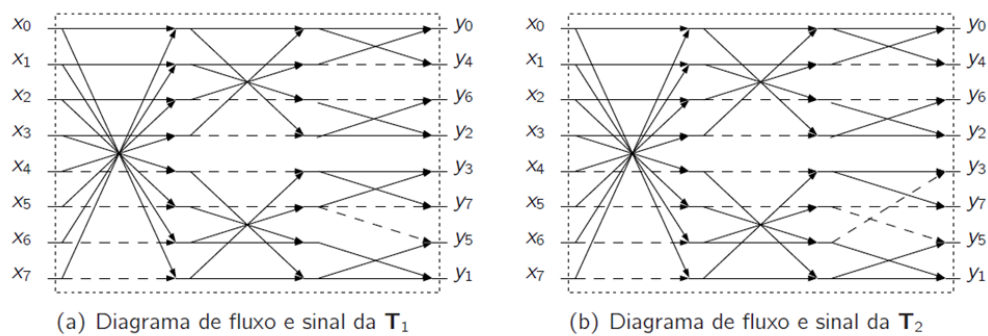


Figura 1

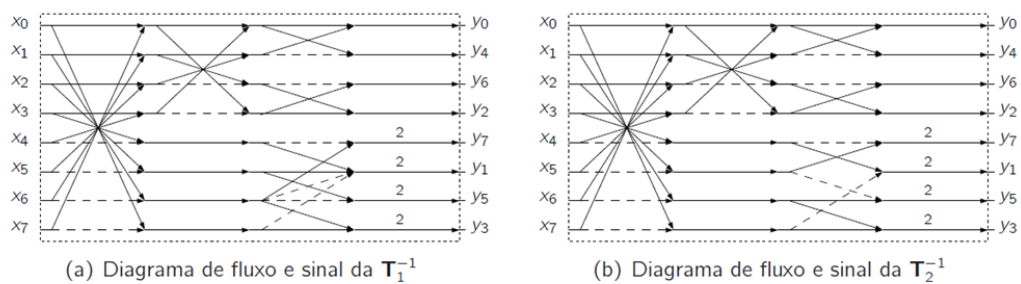


Figura 2

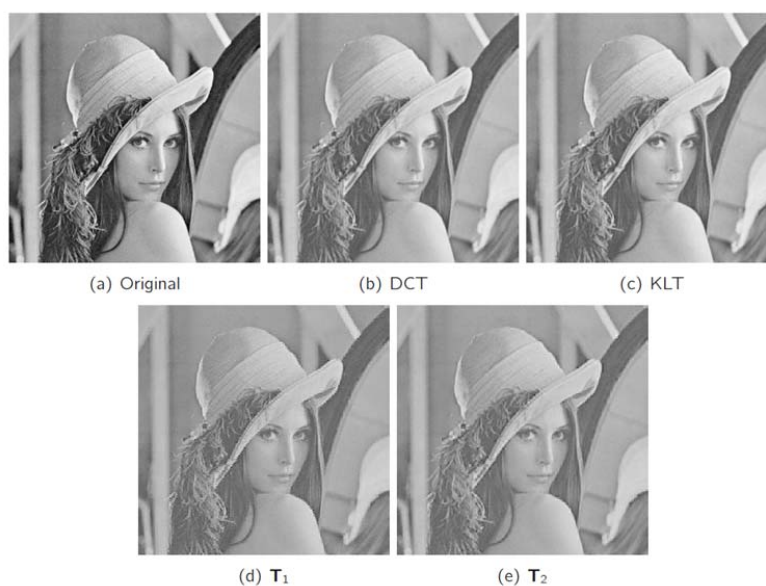


Figura 3

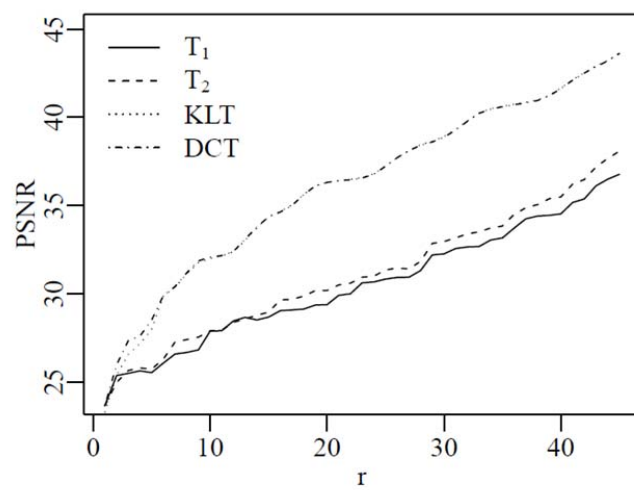


Figura 4

**RESUMO****APROXIMAÇÃO PARA A TRANSFORMADA DE KARHUNEN-LOÈVE  
COM BAIXO CUSTO ARITMÉTICO COM APLICAÇÕES EM  
COMPRESSÃO DE IMAGENS**

A transformada de Karhunen-Loève (KLT) é frequentemente utilizada na descorrelação de dados e na redução de dimensionalidade. Por sua característica de compactação de energia em poucos componentes, encontram-se aplicações importantes da KLT no contexto de compressão de imagens. Contudo, seu uso em aplicações reais é impossibilitado devido ao alto custo computacional requerido para seu cálculo a cada sinal de entrada. Neste contexto, a presente invenção propõe uma versão de baixo custo computacional para a KLT por meio da aplicação da função sinal nos elementos da matriz induzida pela KLT. Denominamos a transformada proposta por KLT sinalizada (SKLT). É dada ênfase nas transformadas de comprimento 8 devido à utilização desse comprimento de transformadas na codificação de imagem, como no padrão JPEG. A SKLT foi numericamente avaliada por meio de medidas de codificação e de distância totais para a KLT exata.