Speckle Filtering based on Stochastic Distances and Tests between Distributions

Leonardo Torres[1]
Alejandro C. Frery[2]

[1] Instituto Nacional de Pesquisas Espaciais – INPE
Caixa Postal 515 - 12227-010 - São José dos Campos – Brazil
ljmtorres@dpi.inpe.br

[2]Laboratório de Computação Científica e Análise Numérica – LaCCAN
Universidade Federal de Alagoas – UFAL
57072-970 - Maceió, AL – Brazil
acfrery@gmail.com

Abstract. This paper presents and evaluates a filter design based on stochastic distances and tests between distributions. A window is defined around each pixel, overlapping samples are compared and only those which pass a goodness-of-fit test are used to compute the filtered value. The technique is applied to intensity SAR data with homogeneous regions using the Gamma model. The proposal is compared with the Improved Lee filter using a protocol based on Monte Carlo. Among the criteria used to quantify the quality of filters, we employ the equivalent number of looks, line and edge preservation. Moreover, we also assessed the filters by the Universal Image Quality Index and the Pearson's correlation on edges regions. Application in the real images are assessed too. The assessment and the proposed method show good results.

Key words: Speckle, SAR data, Stochastic Distances, Information Theory

1. Introduction

Synthetic Aperture Radar (SAR) data are generated by a system of coherent illumination and are affected by the interference coherent of the signal. It is known that these data incorporate a granular noise that degrades its quality, known as speckle noise, which is also present in the laser, ultrasound-B, and sonar imagery (GOODMAN, 1976). The noise makes the segmentation, extraction, analysis and, classification of objects and information in the image hard tasks.

Statistical analysis is essential for dealing with speckled data. It provides comprehensive support for developing procedures for interpreting the data efficiently, and to simulate plausible images. Different statistical distributions are proposed in the literature to describe speckle data. In this paper we use the Gamma distribution to describe the speckle noise, and a constant to characterize the ground truth (GAO, 2010).

Deledalle et al. (2011) analyzed several similarity criteria for data which depart from the Gaussian assumption, viz., the Gamma and Poisson noises. In Deledalle et al. (2009) the same authors extended the NL-means method to speckled imagery using statistical inference in an iterative procedure. The authors derived the weights using the likelihood function of Gaussian and square root of Gamma (termed "Nakagami-Rayleigh") noises. In Deledalle et al. (2010), the authors proposed the use of a nonlocal approach to estimate jointly reflectivity, phase difference and coherence from a pair of co-registered single-look complex SAR images.

Yang and Clausi (2009) proposed a new method for filtering speckle that explores the similarity based on structural patterns, namely structure-preserving speckle reduction (SPSR). In this method the image is logarithmically transformed to convert the multiplicative noise into additive noise. The SPSR filter uses the nonlocal means method with the objective of preserving structures and edges.

Coupé et al. (2009) also used a logarithmic transformation and assume zero-mean Gaussian noise to propose the Optimized Bayesian NL-means with block selection (OBNLM). The OBNLM filter is an optimized version of the filter proposed by Kervrann et al. (2007) which employs a new distance for comparing patches and selecting the most similar pixels.

In this work we employ explicit expressions of stochastic distances between Gamma random variables (NASCIMENTO et al., 2010) for the selection of patches inspired in the Nagao-Matsuyama filter (NAGAO; MATSUYAMA, 1979). These distances are scaled to possess known asymptotic properties which lead to p-values. These p-values are then transformed by a nonlinear function before defining the NL-means filter.

The paper is organized as follows: Section 2 presents the statistical modeling used to describe speckle data. Section 3 describes the new method for filtering speckle. Section 4 presents the metrics for assessing the quality of the filtered images. Sections 5 and 6 present the results and conclusions.

2. The Multiplicative Model

According to Goodman (1976), the multiplicative model can be used to describe SAR data. This model asserts that the intensity observed in each pixel is the outcome of the random variable $Z\colon \Omega \to \mathbb{R}_+$ which, in turn, is the product of two independent random variables: $X\colon \Omega \to \mathbb{R}_+$, that characterizes the backscatter; and $Y\colon \Omega \to \mathbb{R}_+$, which defines the intensity of the speckle noise. The distribution related to the observed intensity $Z = XY$ is completely specified by the distributions proposed for X and Y.

This paper focus is homogeneous regions in intensity images, so the constant $X \sim \lambda > 0$ defines the backscatter, and $Y \sim \Gamma(L, L)$ models the speckle noise by a Gamma distribution (with expected value $\mathbb{E}(Y) = 1$), where L is equivalent number of looks. Thus, it follows that $Z \sim \Gamma\left(L, L/\lambda\right)$ and its density is

$$f_Z(z; L, \lambda) = \frac{L^L}{\lambda^L \Gamma(L)} z^{L-1} \exp\left\{\frac{-Lz}{\lambda}\right\}, \tag{1}$$

where Γ stands for the Gamma function, $L \geq 1$ and $z, \lambda > 0$. We describe different levels of heterogeneity by allowing the number of looks L to vary locally.

The likelihood of $\boldsymbol{z} = (z_1, z_2, \ldots, z_n)$, a random sample of size n from the $\Gamma(L, L/\lambda)$ law, is given by

$$\mathcal{L}(L, \lambda; \boldsymbol{z}) = \left(\frac{L^L}{\lambda^L \Gamma(L)}\right)^n \prod_{j=1}^{n} z_j^{L-1} \exp\left\{-\frac{L z_j}{\lambda}\right\}. \tag{2}$$

Thus, the maximum likelihood estimator for (L, λ), namely $(\widehat{L}, \widehat{\lambda})$, is given by $\widehat{\lambda} = n^{-1} \sum_{j=1}^{n} z_j$ and by the solution of

$$\ln \widehat{L} - \psi^0(\widehat{L}) - \ln \frac{1}{n} \sum_{j=1}^{n} z_j + \frac{1}{n} \sum_{j=1}^{n} \ln z_j = 0, \tag{3}$$

where ψ^0 is the digamma function (TORRES et al., 2012b).

3. Stochastic Distances Filter

The proposed filter is local and nonlinear, initially proposed by Torres et al. (2012a, 2012b). It is based stochastic distances and tests between distributions (NASCIMENTO et al., 2010), obtained from the class of (h, ϕ)-divergences. The proposal employs the neighborhoods defined by Nagao and Matsuyama (1979).

Each filtered pixel has a 5×5 neighborhood, within which nine areas are defined and treated as different samples. Denote $\widehat{\boldsymbol{\theta}}_1$ the estimated parameter in the central 3×3 neighborhood, and

$\big(\widehat{\boldsymbol{\theta}}_2, \ldots, \widehat{\boldsymbol{\theta}}_9\big)$ the estimated parameters in the eight remaining areas. To account for possible departures from the homogeneous model, we estimate $\widehat{\boldsymbol{\theta}}_i = (L_i, \lambda_i)$, $i = \{1, \ldots, 9\}$ by maximum likelihood.

The proposal is based on the use of stochastic distances on small areas within the filtering window. Consider Z_1 and Z_i random variables defined on the same probability space, characterized by the densities $f_{Z_1}(z_1; \boldsymbol{\theta}_1)$ and $f_{Z_i}(z_i; \boldsymbol{\theta}_i)$, respectively, where $\boldsymbol{\theta_1}$ and $\boldsymbol{\theta_i}$ are parameters. Assuming that both densities have the same support $I \subset \mathbb{R}$, the (h, ϕ)-divergence between f_{Z_1} and f_{Z_i} is given by

$$D_\phi^h(Z_1, Z_i) = h\Big(\int_{x \in I} \phi\Big(\frac{f_{Z_1}(x; \boldsymbol{\theta}_1)}{f_{Z_i}(x; \boldsymbol{\theta}_i)} \Big) \, f_{Z_i}(x; \boldsymbol{\theta}_i) \, \mathrm{d}x \Big), \tag{4}$$

where $h\colon (0, \infty) \to [0, \infty)$ is a strictly increasing function with $h(0) = 0$ and $h'(x) > 0$, $\phi\colon (0, \infty) \to [0, \infty)$ is a convex function for all $x \in \mathbb{R}$. Choices of the functions h and ϕ result in several divergences.

Divergences sometimes do not obey the requirements to be considered distances. A simple solution, described in Nascimento et al. (2010), is to define a new measure d_ϕ^h given by

$$d_\phi^h(Z_1, Z_i) = \frac{D_\phi^h(Z_1, Z_i) + D_\phi^h(Z_i, Z_1)}{2}. \tag{5}$$

Distances, in turn, can be conveniently scaled in order to present good statistical properties that make them test statistics (NASCIMENTO et al., 2010):

$$S_\phi^h(\widehat{\boldsymbol{\theta}}_1, \widehat{\boldsymbol{\theta}}_i) = \frac{2mnk}{m+n} \, d_\phi^h(\widehat{\boldsymbol{\theta}}_1, \widehat{\boldsymbol{\theta}}_i), \tag{6}$$

where $\widehat{\boldsymbol{\theta}}_1$ e $\widehat{\boldsymbol{\theta}}_i$ are maximum likelihood estimators based on samples size m and n, respectively, and $k = \big(h'(0)\phi''(1)\big)^{-1}$. The null hypothesis $\boldsymbol{\theta_1} = \boldsymbol{\theta_i}$ is rejected at a level η, if $\Pr(S_\phi^h > \eta)$, and since under mild conditions S_ϕ^h is χ_M^2 asymptotically distributed, being M the dimension of $\boldsymbol{\theta_1}$, the test is well defined. Details can be seen in the work by Salicrú et al. (1994). The statistical test derived in this paper was the Kullback-Leibler test:

$$S_{KL} = \frac{mn\big(\widehat{L}_1 + \widehat{L}_i\big)}{m+n} \left(\frac{\widehat{\lambda}_1^2 + \widehat{\lambda}_i^2}{2\widehat{\lambda}_1\widehat{\lambda}_i} - 1 \right). \tag{7}$$

The filtering procedure consists in checking which regions can be considered as coming from the same distribution that produced the data which comprises the central block. The sets which are not rejected are used to compute a local mean. If all the sets are rejected, the filtered value is updated with the average on the 3×3 neighborhood around the filtered pixel.

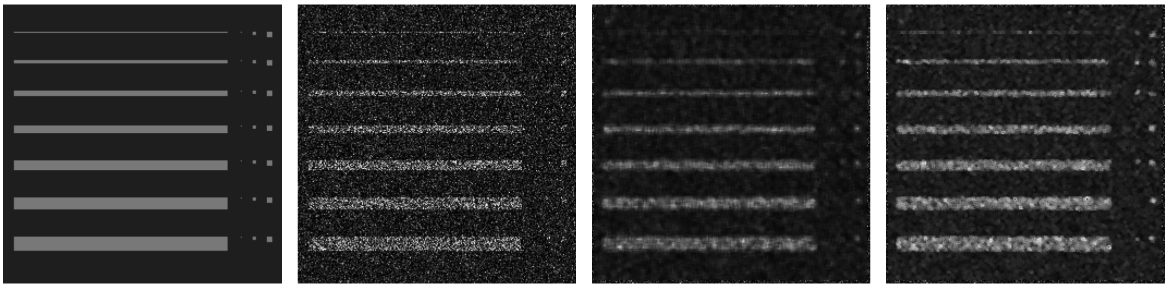
4. Image Quality Assessment

Image quality assessment in general, and filter performance evaluation in particular, are hard tasks (MOSCHETTI et al., 2006; WANG; BOVIK, 2002). Moschetti et al. (2006) discussed the need of making a Monte Carlo study when assessing the performance of image filters. They proposed a protocol which consists of using a phantom image (see Figure 1(a)) corrupted by speckle noise (see Figure 1(e)). The experiment consists of simulating corrupted images as matrices of independent samples of some distribution with different parameters. Every simulated image is subjected to filters, and the results are compared (see Figures 1(f) and 1(g)).

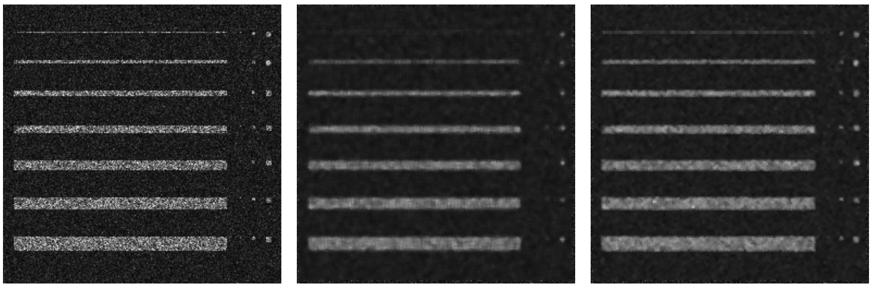
Among the criteria used to quantify the quality of the filters, we employ (MOSCHETTI et al., 2006):

- **Equivalent Number of Looks**: in intensity imagery and homogeneous areas, it can be estimated by $\mathsf{NEL} = (\bar{z}/\widehat{\sigma}_Z)^2$, i.e., the square of the reciprocal of the coefficient of variation. In this case, the bigger the better.
- **Line Contrast**: the preservation of the line of one pixel width will be assessed by computing three means: in the coordinates of the original line (x_ℓ) and in two lines around it (x_{ℓ_1} and x_{ℓ_2}). The contrast is then defined as $2x_\ell - (x_{\ell_1} + x_{\ell_2})$, and compared with the contrast in the phantom. The best values are the smallest.
- **Edge Preserving**: it is measured by means of the edge gradient (the absolute difference of the means of strip around edges) and variance (same as the former but using variances instead of means). The best values are the smallest.

A "good" technique must combat speckle and, at the same time, preserve details as well as relevant information.



(a) Phantom (b) Corrupted, 1-look (c) Improved Lee filter (d) Kullback-Leibler filter



(e) Corrupted, 4-looks (f) Improved Lee filter (g) Kullback-Leibler filter

Figure 1. Lee Protocol phantom, speckled data and filtered images.

Furthermore, we also assessed the filters by the universal image quality index (WANG; BOVIK, 2002) and the correlation measure β_ρ. The universal image quality index is defined by

$$Q = \frac{s_{xy}}{s_x s_y} \frac{2\overline{x}\overline{y}}{\overline{x}^2 + \overline{y}^2} \frac{2 s_x s_y}{s_x^2 + s_y^2}, \tag{8}$$

where $s^2_\bullet$ and $\overline{\bullet}$ denote the sample variance and mean, respectively. The range of Q is $[-1, 1]$, being 1 the best value. The quantity

$$\beta_\rho = \frac{\sum_{j=1}^n (x_j - \bar{x})(y_j - \bar{y})}{\sqrt{\sum_{j=1}^n (x_j - \bar{x})^2 \sum_{j=1}^n (y_j - \bar{y})^2}}, \tag{9}$$

is a correlation measure is between the Laplacians of images X and Y, where $\bullet_j$ and $\overline{\bullet}$ denote the gradient values of the jth pixel and mean of the images $\nabla^2 X$ and $\nabla^2 Y$, respectively. The range of β_ρ is $[-1, 1]$, being 1 perfect correlation.

The Figure 2 shows a block diagram for the method of assessment of the proposal.

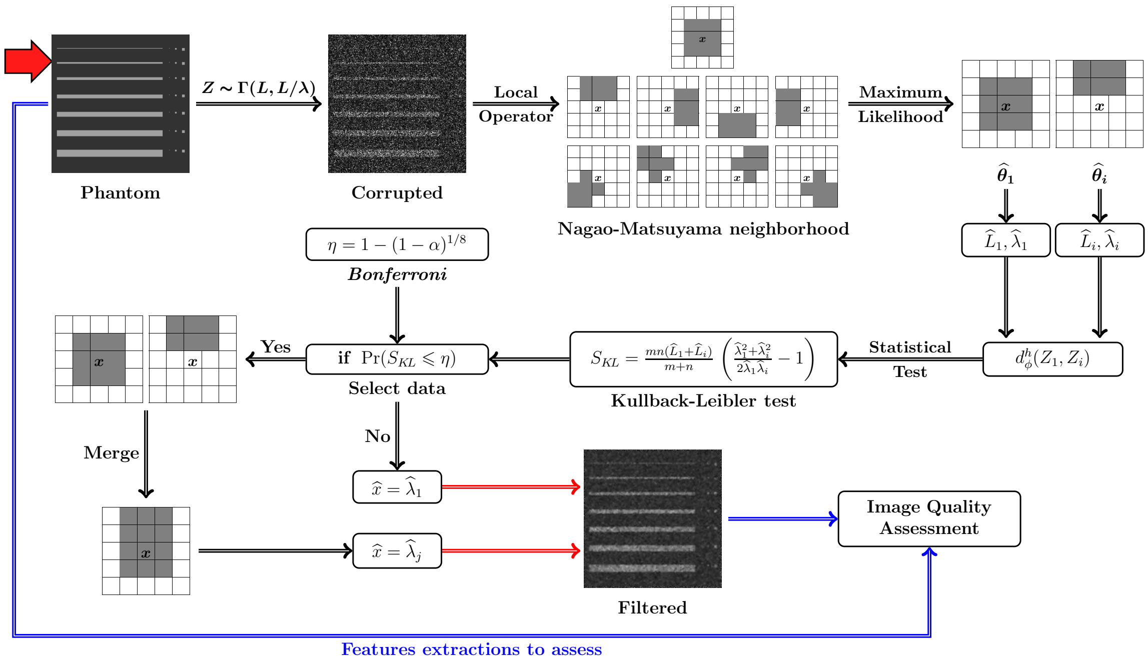


Figure 2. Block diagram for assessment of the proposed technique.

5. Results and Analysis

The proposal, unlike Torres et al. (2012a, 2012b), was compared with the Improved Lee filter (LEE et al., 2009). The tests were performed at the 95% level of significance. The results obtained are summarized by means of boxplots (see Table 1). Each boxplot describes the results of one filter, generating 100 independent $L = \{1, 4\}$ looks images, mean background $\lambda = 30$ and mean lines $\lambda = 120$.

The best values are highlighted in bold in Table 1. In the first and second situation, the Improved Lee filter provides better results by a very small margin regarding the equivalent number of looks and, by a large margin, with respect to edge variation only on first situation. In all cases the differences are significative. This may be the result of the heavier blurring introduced by this filter, as shown later in the application to real images.

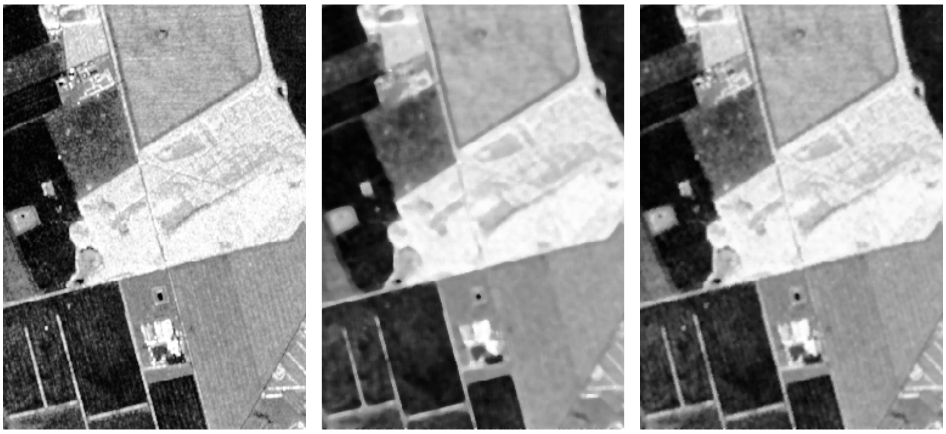
5.1 Real data

Not every quality measure presented in Section 4 can be applied to real data, unless the ground truth is known. One of the quality measures that can be used in this case is the Universal Image Quality Index by Wang and Bovik (2002).

Figure 3 presents the results of applying the filters to an image obtained by the Danish EMISAR L-band fully polarimetric sensor over agricultural fields of Foulum. The original 250×350 pixels image of the HH intensity band is shown in Figure 3(a), its filtered versions by the Improved Lee and Kullback-Leibler filters are presented in Figures 3(b) and 3(c), respectively.

6. Conclusions

This paper presented an assessment of the filter based on stochastic distances for speckle noise reduction. The proposal was compared with the Improved Lee filter, using a protocol based on Monte Carlo experiences. Moreover, the β_ρ and Q index were used to assert the proposal. The proposed filters behave alike, and they outperform the Improved Lee filter in five



(a) SAR data (HH polarization) (b) Improved Lee filter, $Q = 0.540$ (c) Kullback-Leibler filter, $Q = 0.707$

Figure 3. Real data Danish EMISAR L-band (HH polarization) and filtered versions.

out of six quality measures. Other significance levels will be tested, along with different points of the parameter space in order to have a more complete assessment of the proposal.

References

COUPÉ, P.; HELLIER, P.; KERVRANN, C.; BARILLOT, C. Nonlocal means-based speckle filtering for ultrasound images. **IEEE Transactions Image Processing**, v. 18, n. 10, p. 2221–2229, 2009.

DELEDALLE, C.-A.; DENIS, L.; TUPIN, F. Iterative weighted maximum likelihood denoising with probabilistic patch-based weights. **IEEE Transactions on Image Processing**, v. 18, n. 2, p. 2661–2672, 2009.

DELEDALLE, C.-A.; TUPIN, F.; DENIS, L. A non-local approach for SAR and interferometric SAR denoising. In: **IEEE International Geoscience and Remote Sensing Symposium (IGARSS)**. Honolulu: [s.n.], 2010. p. 714–717.

DELEDALLE, C.-A.; TUPIN, F.; DENIS, L. Patch similarity under non gaussian noise. In: **IEEE International Conference on Image Processing (ICIP)**. [S.l.: s.n.], 2011. p. 1845–1848. ISSN 1522-4880.

GAO, G. Statistical modeling of SAR images: A Survey. **Sensors**, v. 10, n. 1, p. 775–795, 2010.

GOODMAN, J. W. Some fundamental properties of speckle. **Journal of the Optical Society of America**, v. 66, n. 11, p. 1145–1150, 1976.

KERVRANN, C.; BOULANGER, J.; COUPÉ, P. Bayesian non-local means filter, image redundancy and adaptive dictionaries for noise removal. In: **Proc. Conf. Scale-Space and Variational Meth**. Ischia: [s.n.], 2007. p. 520–532.

LEE, J.-S.; WEN, J.-H.; AINSWORTH, T. L.; CHEN, K.-S.; CHEN, A. J. Improved sigma filter for speckle filtering of SAR Imagery. **IEEE Transactions on Geoscience and Remote Sensing**, v. 47, n. 1, p. 202–213, 2009. ISSN 0196-2892.

MOSCHETTI, E.; PALACIO, M. G.; PICCO, M.; BUSTOS, O. H.; FRERY, A. C. On the use of Lee's protocol for speckle-reducing techniques. **Latin American Applied Research**, v. 36, n. 2, p. 115–121, 2006.

NAGAO, M.; MATSUYAMA, T. Edge preserving smoothing. **Computer Graphics and Image Processing**, v. 9, n. 4, p. 394–407, 1979.

NASCIMENTO, A. D. C.; CINTRA, R. J.; FRERY, A. C. Hypothesis testing in speckled data with stochastic distances. **IEEE Transactions on Geoscience and Remote Sensing**, v. 48, n. 1, p. 373–385, 2010.

SALICRÚ, M.; MORALES, D.; MENÉNDEZ, M. L.; PARDO, L. On the applications of divergence type measures in testing statistical hypotheses. **Journal of Multivariate Analysis**, v. 21, n. 2, p. 372–391, 1994.

TORRES, L.; CAVALCANTE, T.; FRERY, A. C. A new algorithm of speckle filtering using stochastic distances. In: **IEEE International Geoscience and Remote Sensing Symposium (IGARSS)**. Munich: [s.n.], 2012. Available from: <http://arxiv.org/abs/1308.6487>.

TORRES, L.; CAVALCANTE, T.; FRERY, A. C. Speckle reduction using stochastic distances. In: L. Alvarez et al. (Ed.). **Pattern Recognition, Image Analysis, Computer Vision, and Applications**. Buenos Aires: Springer, 2012. (Lecture Notes in Computer Science, v. 7441), p. 632–639.

WANG, Z.; BOVIK, A. C. A universal image quality index. **IEEE Signal Processing Letters**, v. 9, n. 3, p. 81–84, 2002.

YANG, X.; CLAUSI, D. A. Structure-preserving speckle reduction of SAR images using nonlocal means filters. In: **IEEE International Conference on Image Processing**. Cairo: [s.n.], 2009. p. 2985–2988.

Table 1. Statistics of the metrics of simulated images: 100 replications with 1-iteration.

	Speckle	SAR Measures				Q Index	β_ρ Index
	Filter	NEL	Line Cont.	Edge Grad.	Edge Var.		
1-look	Lee	**14.393** (1.32)	1.877 (0.02)	76.122 (5.94)	**0.762** (0.65)	0.151 (0.002)	0.751 (0.007)
	KL	11.305 (1.09)	**1.831** (0.03)	**69.447** (7.70)	4.690 (1.59)	**0.216** (0.002)	**0.817** (0.010)
4-looks	Lee	**95.494** (13.41)	1.798 (0.02)	59.415 (5.09)	8.065 (1.27)	0.218 (0.001)	0.842 (0.003)
	KL	75.206 (9.67)	**1.758** (0.05)	**47.114** (5.72)	**6.830** (2.31)	**0.263** (0.001)	**0.895** (0.007)