

Systematic Error Analysis for the Enhancement of Biometric Systems Using Soft Biometrics

Anastasios Drosou, Dimitrios Tzovaras, *Member, IEEE*, Konstantinos Moustakas, *Member, IEEE*, and Maria Petrou, *Senior Member, IEEE*

Abstract—This letter presents a novel probabilistic framework for augmenting the recognition performance of biometric systems with information from continuous soft biometric (SB) traits. In particular, by modelling the systematic error induced by the estimation of the SB traits, a modified efficient recognition probability can be extracted including information related both to the hard and SB traits. The proposed approach is applied without loss of generality in the case of gait recognition, where two state-of-the-art gait recognition systems are considered as hard biometrics and the height and stride length of the individuals are considered as SBs. Experimental validation on two known, large datasets illustrates significant advances in the recognition performance with respect to both identification and authentication rates.

Index Terms—Feature space partitioning, gait recognition, soft biometrics, systematic error.

I. INTRODUCTION

IT is a common place that security in computer systems is an increasingly critical issue that affects a series of diverse applications, ranging from granting access control in restricted infrastructures to e-commerce transactions. Such applications require reliable personal recognition schemes to either confirm or determine the identity of an individual requesting their services. To this extent, biometrics have been proven to provide unique and powerful advantages over other traditional technologies for identity verification, such as PINs or tokens that can be easily forgotten, lost or stolen.

In general, biometric traits can be divided in two main categories. Namely, **hard biometrics** include both the common physical biometrics (e.g., fingerprint) and the behavioural ones that describe activity-related patterns of the user (e.g., gait [1]), while **soft biometrics** can be divided into continuous (e.g., height, weight, etc.) or discrete human characteristics (i.e., gender, race/ethnicity etc.) [2].

Soft biometrics (SBs) lack the distinctiveness and permanence to sufficiently differentiate any two individuals and thus, they cannot deterministically predict the identity of the user [4].

Manuscript received June 19, 2012; revised September 17, 2012; accepted September 18, 2012. Date of publication October 02, 2012; date of current version October 10, 2012. This work was supported by the Greek national research project BIOTAFTOTITA. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Mujdat Cetin.

A. Drosou and M. Petrou are with the Department of Electrical and Electronic Engineering, Imperial College London, SW7 2AZ, London, U.K. (e-mail: a.drosou09@imperial.ac.uk).

D. Tzovaras and M. Petrou are with Information Technologies Institute, Centre for Research and Technology Hellas, 57001 Thessaloniki, Greece (e-mail: tzovaras@iti.gr; petrou@iti.gr).

K. Moustakas is with the Electrical and Computer Engineering Department, University of Patras, Patra, Greece e-mail: moustakas@ece.upatras.gr.

Digital Object Identifier 10.1109/LSP.2012.2221701

However, they manage to straightforwardly provide useful information towards user identification in large datasets [3], by either verifying hypotheses or reducing the search space in typical biometric systems [6]. Specifically, they allow the overall query to be answered better and faster by contributing to a much smaller candidate pool and by minimizing the size of comparisons. Moreover, it has been stated that the automatic inference of SB traits outperforms human observation, which may be subconsciously biased [7].

The idea of improving the performance of typical biometric systems via combination with soft traits has been proposed in the bibliography several times until today. For instance, Nandakumar *et al.* proposed quality-based fusion of static biometrics for the improvement of the recognition accuracy [8]. Later, Moustakas *et al.* have proposed a Bayesian framework for improving the authentication rates from behavioral gait traits with gait-related SBs [9], while Marcialis *et al.* proposed a similar approach for facial recognition in [10].

A. Contribution

The current letter proposes a novel and highly efficient probabilistic framework for augmenting the recognition performance in biometric systems via the integration of one or more SB traits and the modelling of the induced systematic error during their measurement.

This way, some serious open issues from [9] and [10] are addressed. In [10], the users are categorized into “minority” and “majority” groups, according to the frequency of appearance of their SB traits. Moreover, only the extreme cases of SB traits are boosted in [9], while the fact that they are defined in a single dimension leads to a uniform, linear quantization of the feature space, not applicable for most real scenarios. Last but not least, despite its seemingly smooth function, both works are based on the invalid assumption of independent conditional probabilities of the geometric trait when multiple SBs are available.

In this context, the following framework proposes a two-step process for augmenting the matching scores of genuine users, initially by assigning them to specific clusters in the feature space and then by evaluating the systematic error of the incoming measurement according to a predefined statistical model. Most important, the error measurements of the SBs are claimed to be independent, since they are in direct connection with distinct measurement processes.

II. AUGMENTING RECOGNITION PERFORMANCE VIA SOFT BIOMETRICS

Let Ω be the set of all identities in the M -sized user population $\Omega = \{\omega_1, \omega_2, \dots, \omega_M\}$, x_c be the hard biometric information (e.g., geometric gait) and x_s be a continuous soft biometric (SB) trait (i.e., the height of the user) from a set X with N available SBs $X = \{x_{s_1}, x_{s_2}, \dots, x_{s_N}\}$. As such, $p(\omega|x_c) =$

$1 - p(\bar{\omega}|x_c)$ is the matching score of the conventional biometric system.

A. Partitioning the Feature Space

Contrary to the simple 3-stage partitioning (i.e., small-, normal- and large-sized population) [9], [10], a more sophisticated spatial partitioning of the feature space F in N_C clusters C_i , that exhibit notable variation in terms of their defining SBs, is proposed herein.

The authentication probability of a client user is augmented, when the incoming SB traits refer to the same cluster as the claimed ID. In all other cases, the matching probability $p(\omega|x_c)$ remains intact and is solely based hard biometric trait of the user. Although there is no actual limitation in the dimensionality of the feature space, the simple case of $2D$ clusters will be studied herein, without loss of generality.

In this respect, a cluster C_i of the multidimensional soft biometric feature space, associated to a subset $S_i \in \Omega$ of the set of identities Ω , is characterized as a valid cluster iff the following hold.

The a-priori probability of an identity ω to belong to a cluster has to be low:

$$0 < p(\omega \in C_i | x_{s_1}, \dots, x_{s_N}) = p_i(\omega) \ll 1$$

and there should exist a subset S_i of Ω , so as

$$\exists S_i \subset \Omega \begin{cases} \forall \omega \in S_i, p(x_s \in C_i | \omega) > \alpha \\ \forall \omega \notin S_i, p(x_s \in C_i | \omega) \approx p_i(\omega) \end{cases}$$

where α is a minimum non-zero value, C is the union of all clusters whose number N_C should be significantly lower than the size $|\Omega|$ of the identity set Ω :

$$C = \cup C_i, \forall i = 1, \dots, N_C$$

$$N_C \ll |\Omega|$$

Three different partitioning alternatives have been implemented herein, all of which fulfil the requirements of a cluster:

Uniform Orthogonal Clustering (UOC) is a linear way of partitioning the feature space. The dimensions of the prototype orthogonal cluster are defined by using a brute force iterative algorithm on an adequately large reference SB feature dataset. The major drawback of the current clustering method, is that it deals with each biometric feature separately. This way, some clusters are left “empty”, while some other are “overcrowded”.

Hexagonal Cell Clustering (HCC) is a more efficient way for partitioning the feature space. The isotropy is preserved along the whole feature space via data normalization to the standard deviation, while a more uniform distribution of the samples among the cells is provided. Hereby, only the radius of the prototype hexagon has to be estimated.

Gaussian Clustering (GC) refers to the generation of multidimensional gaussian clusters on the feature space and offers increased flexibility in grouping similar users. To this direction, an unsupervised clustering approach was implemented, based on the ISODATA and the expectation-maximization (EM) algorithms, so that the whole feature space is described as mixture of multidimensional Gaussians

$$\mathcal{N}(\mathbf{f}_\omega | \mu_k, \Sigma_k) = \frac{1}{(2\pi)^{Z/2} |\Sigma_k|^{1/2}} e^{-\frac{1}{2}(\mathbf{f}_\omega - \mu_k)^T \Sigma_k^{-1} (\mathbf{f}_\omega - \mu_k)}$$

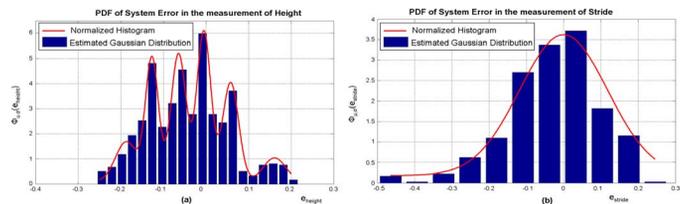


Fig. 1. Distribution of the Systematic Error in (a) Height and (b) Stride Measurements and the corresponding fitting curve.

Vector \mathbf{f}_ω includes all utilized SB trait values, $\mathbf{f}_\omega = \{x_{s_{n,1}}(\omega), \dots, x_{s_{n,Z}}(\omega)\}$, while μ_k and Σ_k are the Z -dimensional mean vector and the $Z \times Z$ covariance matrix of the k th Gaussian, respectively.

At the authentication stage, the assignment of an incoming SB feature vector of a user to a cluster is performed according to the maximum likelihood (ML) criterion.

B. Modelling the Systematic Error Probability

Let us now define the ground truth value $x_{s_n}^g$ as the SB trait n of user ω and $\tilde{x}_{s_n,l}$ as the l th value measured by the system ($\tilde{X}_{s_n}(\omega) = \{\tilde{x}_{s_n,1}(\omega), \dots, \tilde{x}_{s_n,L}(\omega)\}$, where L is the total number of measurements. For an adequately large number $T = M \times L$ of measurements, the noise distribution that is induced as error in the measurement (i.e., noise) by the system can be estimated as described hereafter.

As long as T is large enough for reliable statistical estimates, the normalized values $e_{s_n,l}(\omega_m) = \tilde{x}_{s_n,l}(\omega_m) - x_{s_n}^g(\omega_m)$ can be produced. Having these data for the whole registered population, it is trivial to fit the normalized values distribution by a $1D$ Gaussian Mixture of the following type:

$$p(e_s | \omega) = \sum_{k=1}^K \pi_k \mathcal{N}_p(e_s | \mu_k, \sigma_k)$$

where $\mathcal{N}_p(e_s | \mu_k, \sigma_k)$ stands for the k th single Gaussian distribution that contributes to the mixture. The values π_k, μ_k and σ_k can be easily computed by utilizing the iterative Expectation-Maximization (EM) algorithm on the data’s histogram, until convergence. The initial parameter regarding the number K of single Gaussian distributions in the $1D$ mixture model is experimentally selected, as the one that produces an acceptable error value in the χ^2 - test.

Consequently, $p(e_s | \bar{\omega})$ can be calculated as

$$p(e_s | \bar{\omega}) = \frac{p(e_s) - p(\omega)p(e_s | \omega)}{1 - p(\omega)} \quad (1)$$

where $p(\omega) = 1/M$ and $p(e_s) = 1/L$ are priors.

At this point, it becomes evident that a separate clustering process in the multidimensional measurement error space would be meaningless, since the feature distribution is known, from the corresponding probability density functions (pdf) of the measurement errors are known (Fig. 1) and that the measurement errors themselves are i.i.d. variables. The exploitation of the information provided by the raw soft biometric features is herein achieved, as described in Section II.A, so as to further improve the recognition performance.

C. Estimation of the Recognition Probability

In this paragraph, the augmentation process is presented that is applied only to these users, whose SB traits resemble the claimed ones. Yet, it is important to highlight that the previous

frameworks for augmenting biometric recognition with SBs assumed independence between the SBs in an ad-hoc manner, which does not hold per se. Herein, the independence between the inserted systematic error for each SB is guaranteed by definition, since the produced distributions refer to the uncorrelated measurement errors (not to the SB traits).

In this context, the goal herein is to find a generic expression of the conditional probability $p_{\omega}^{e_s} = p(\omega | x_c, e_{s_1}, \dots, e_{s_N})$ that denotes the final recognition score:

$$p(\bar{\omega} | x_c, e_{s_1}, \dots, e_{s_N}) = 1 - p(\omega | x_c, e_{s_1}, \dots, e_{s_N}) \quad (2)$$

while according to Bayes' theorem

$$\bar{p}_{\omega}^{e_s} = p(\bar{\omega} | x_c, e_{s_1}, \dots, e_{s_N}) = \frac{p(x_c, e_{s_1}, \dots, e_{s_N} | \bar{\omega}) p(\bar{\omega})}{p(x_c, e_{s_1}, \dots, e_{s_N})}$$

By assuming that the error measurements stem from distinct and uncorrelated measurement processes, the variables e_{s_n} can be held as *i.i.d.* As such, $p(x_c, e_{s_1}, \dots, e_{s_N}) = p(x_c) p(e_{s_1}) \dots p(e_{s_N})$.

Finally, according to the calculations presented in <http://www.iti.gr/~drosou/SystematicError/EquationProof.pdf>, (2) is expressed by the following generic formula

$$p_{\omega}^{e_s} = 1 - \bar{p}_{\omega}^{e_s} = 1 - \prod_{n=1}^N p(e_{s_n} | \bar{\omega}) p(\bar{\omega} | x_c)$$

III. CASE STUDY: ENHANCING GAIT RECOGNITION WITH SOFT BIOMETRICS

Gait recognition is offering high level of unobtrusiveness and performance. Contrary to model-based approaches, feature-based methods are generally seen to perform better in terms of both accuracy and computational complexity, while other limitations, such as clothing or slight changes in walking direction, are compensated via non-canonical view gait recognition and superfluous clothing removal [5].

Herein, the feature vector x_c (see Section II), that refers to the dynamic gait features (i.e., hard biometric), is extracted using two gait recognition algorithms. The first algorithm is presented in [11] and is based on the two well know Radon Transforms that are applied to gait sequence silhouettes (i.e., *BS - RIT* and *BS - CIT*). The second algorithm is based on matching spatiotemporal descriptors of the human gait, the so-called Gait Energy Images (i.e., *GEI - RIT* and *GEI - CIT*).

Moreover, the “height” and “stride length” soft biometric (SB) features should be extracted. This is trivially achieved from the stereoscopic gait sequences, as the highest-lowest part of the subject and to the largest distance between the legs within a gait cycle, respectively. It becomes evident that the process followed for the estimation of the stride length is prone to bad illumination and the corresponding shadows that are created on the walking floor, which may have as result the occlusion of the edges of the feet. On the other hand, since the height is estimated as the mean height value of all recorded frames is robust to illumination changes along the walking path. Thus, the errors in measurements mainly stem from possible variations in the types of shoes/hills worn by the users and from the natural hopping of humans during walking. Further, uncorrelatedness between the error measurements between the soft biometric traits is supported, according to the correlation

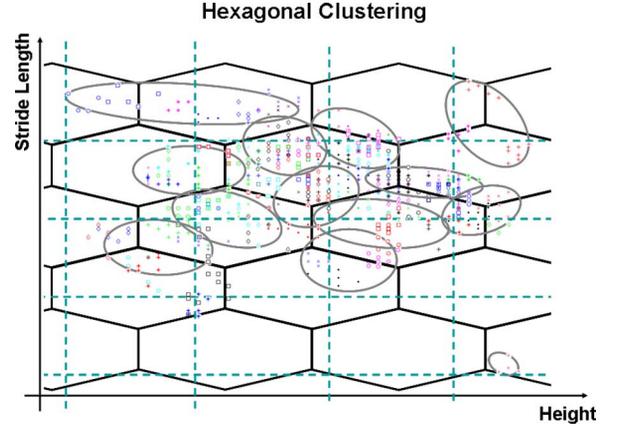


Fig. 2. Three alternatives for partitioning the feature space are studied: (a) UOC, (b) UHC and (c) non-linear 2D GC.

analysis performed on our data (<http://www.iti.gr/~drosou/SystematicError/SBsIID.pdf>). In this context, the independent probabilities $p(e_{s_n} | \omega)$ are modelled as noise-related *pdfs*.

Following the steps in Section II.C, the parameters of the Gaussian Mixtures that best fit the normalized height and stride measurement's errors are illustrated in Fig. 1(a) and Fig. 1(b), respectively (see (1)).

Finally, the three described methodologies (i.e., UOC, HCC, GC) for partitioning of the feature space are applied and the outcomes are illustrated in Fig. 2. In the following Section intensive tests of the proposed methodology are performed on two well known and large 3D-gait databases, so as to evaluate the integrity and the recognition performance of the system.

IV. EXPERIMENTAL RESULTS

The proposed framework was tested both in terms of state of the art curves (i.e., ROC, CMS and score distributions) and experimental evaluation on well known datasets, whereby sequences from different recording sessions are used for enrolment (“gallery”) and identification/authentication (“probe”).

The proposed algorithms have been tested in both the HUMABIO and ACTIBIO databases that include gait sequences captured with stereoscopic cameras. The HUMABIO database was captured in an indoor environment and includes two sessions of 75 and 51 subjects [11], respectively, while the ACTIBIO database includes multiple repetitions on similar scenarios in two sessions with 28 subjects, covering a wide range of real-world use-cases. These datasets (esp. the HUMABIO dataset) are adequately large, compared with other SoA datasets for gait recognition, to extract safe conclusions regarding the performance.

Herein, only the gallery measurements of the HUMABIO database have been used as the reference for both error modelling (Fig. 1) and feature space partitioning (Fig. 2). The performance of the system was evaluated via the probe recordings. Similarly, the ground truth values for the SB data of each user have been measured by a manual annotator on the recorded 3D data. Moreover, these measurements have been verified via actual (i.e., real world) measurements and questionnaires during the capturing of the databases.

Concerning the authentication performance of the proposed approach, the False Acceptance (FAR) and False Rejection Rates (FRR) are illustrated in Fig. 3(a) and 3(b) for the GEI-RIT experiment on both databases. It should be highlighted that the proposed framework manages to decrease the FAR and FRR

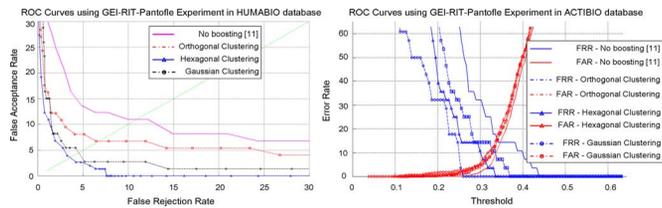


Fig. 3. Receiver Operating Characteristics (ROC) for the GEI-RIT algorithm.

TABLE I
EER SCORES COMPARISON BETWEEN THE PROPOSED METHOD AND [9]

Experiment	Initial	[9]	UOC	HCC	GC
ACTIBIO - BS/RIT	28%	16%	11%	8.5%	4.3%
HUMABIO - BS/RIT	19%	16%	13.5%	9%	3.25%
ACTIBIO - GEI/RIT-Time	25%	15%	18.5%	14.8%	11.1%
ACTIBIO - BS/RIT-Time	28%	15.2%	17.3%	14.8%	11.5%
HUMABIO - BS/CIT-Time	17.5%	15%	8.07%	6.9%	5.2%

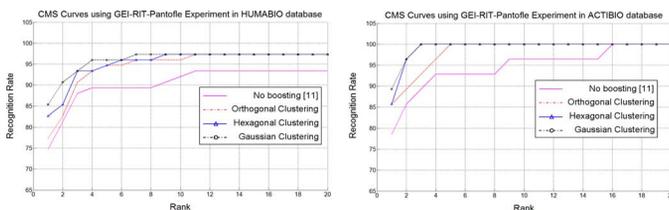


Fig. 4. Cumulative Matching Scores (CMS) for the GEI-RIT algorithm.

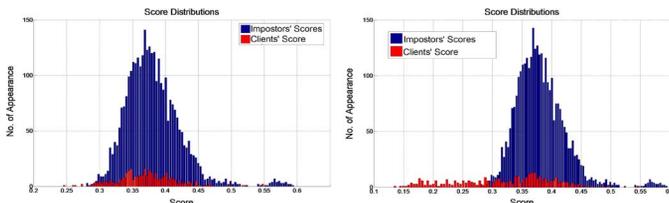


Fig. 5. Scores Distribution (a) before and (b) after the application of the proposed framework (GC) in the HUMABIO RIT-Time experiment.

in the equal error rate EER point from 15.28% to 3.57% in the (GEI-RIT experiment) ACTIBIO database in the GC case, while slightly lower improvements can be noticed in the OC and the HC cases.

Moreover, it is notable that the proposed approach significantly improves the behaviour of the False Rejection Rate of the system (Fig. 3(b)), while only slight improvements can be noticed in its False Acceptance performance. This is achieved by boosting the matching score of the client users above the rejection threshold value. Moreover, both Fig. 3 and Table I exhibit the fact that the proposed system performs equally good in both small datasets (i.e., ACTIBIO), but also in demanding cases of large datasets (i.e., HUMABIO).

The ROC diagram representation in Fig. 3(a) was chosen so that they are consistent and straightforwardly comparable with the results in [11], while the representation in 3(b) was selected, so as to provide the optimal threshold value.

An identification experiment includes the set of all authentication experiments for all subjects of the database. In particular, during an identification experiment, there is no specific claimed ID and thus, the incoming subject is compared with all stored IDs in the database separately. In this respect, Fig. 4 presents the comparative identification results on both datasets for the

GEI-RIT experiments. In particular, four curves are displayed in each figure, which correspond to the Cumulative Matching Scores (CMS), using (a) solely the dynamic gait features and (b) the proposed clustering techniques prior to the application of the attenuation algorithm.

As expected, the more precise the partitioning of the SB feature space, the higher the performance. The most appropriate authentication threshold is defined by the value corresponding to the cross-section of FAR and FRR curves in Fig. 3(b), while in an identification test, the user is assigned the best matching ID, if it exceeds the aforementioned threshold.

Equally significant improvements in the recognition performance further noted in the experiments, shown in Table I. Thereby, a quantitative comparison with the framework, proposed in [9], is attached. The performance improvements become even more notable in difficult scenarios, such as the Time-Scenarios [11], where the users are authenticated 6 months after their enrollment. A clear separation between impostors and clients scores before (Fig. 5(a)) and after (Fig. 5(b)) applying the proposed methodology are presented.

V. CONCLUSION

A novel probabilistic framework for augmenting biometric recognition algorithms via soft biometrics (SB) was proposed. Hereby, the SB-related partitioning of the feature space and the probabilistic modelling of the independent systematic error during soft biometric measurements are seamlessly combined with gait biometrics, so that fusion at score level is avoided. Experimental validation in biometric recognition proved significant improvements in efficiency, authentication and identification potential. An extension to a SB feature space of higher dimensionality is expected to further improve performance. The current framework can be directly applied to any biometric system detecting at least one SB trait.

REFERENCES

- [1] M. S. Nixon, T. Tan, and R. Chellappa, *Human Identification Based on Gait, International Series on Biometrics*, ser. International Series on Biometrics. Berlin, Germany: Springer, 2006.
- [2] A. Dantcheva, C. Velardo, A. D'Angelo, and J.-L. Dugelay, "Bag of soft biometrics for person identification," *Multimedia Tools Applicat.*, vol. 51, pp. 739–777, Oct. 2010.
- [3] A. Dantcheva, J.-L. Dugelay, and P. Elia, "Soft biometrics systems: Reliability and asymptotic bounds," in *4th IEEE Int. Conf. Biometrics: Theory, Applications and Systems (BTAS)*, 2010, pp. 1–6.
- [4] A. K. Jain, S. C. Dass, and K. Nandakumar, "Can soft biometric traits assist user recognition?," in *Proc. SPIE*, 2004, pp. 561–572.
- [5] T. S. Zheng, J. Zhang, K. Huang, and R. He, "Robust view transformation model for gait recognition," in *18th IEEE Int. Conf. Image Processing*, 2011, pp. 2073–2076.
- [6] J. L. Wayman, "Large-scale civilian biometric systems - issues and feasibility," in *Proc. Card Tech/Secur Tech ID*, 1997, pp. 137–156, ver. 1.2.
- [7] K. Ricanek and B. Barbour, "What are soft biometrics and how can they be used?," *Computer*, vol. 44, no. 9, pp. 106–108, 2011.
- [8] K. Nandakumar, Y. Chen, and A. K. Jain, "Quality-based score level fusion in multimetric systems," in *Int. Conf. Pattern Recognition (ICPR)*, 2006, pp. 2–5.
- [9] K. Moustakas, D. Tzovaras, and G. Stavropoulos, "Gait recognition using geometric features and soft biometrics," *IEEE Signal Process. Lett.*, vol. 17, no. 4, pp. 367–370, 2010.
- [10] G. Marcialis, F. Roli, and D. Muntoni, "Group-specific face verification using soft biometrics," *J. Vis. Lang. Comput.*, vol. 20, pp. 101–109, 2009.
- [11] D. Ioannidis, D. Tzovaras, I. G. Damousis, S. Argyropoulos, and K. Moustakas, "Gait recognition using compact feature extraction transforms and depth information," *IEEE Trans. Inf. Forensics Secur.*, vol. 2, no. 3, pp. 623–630, 2007.