

Enhancing 3D Face Recognition by a Robust Version of ICP Based on the Three Polar Representation

Amal Rihani^(✉), Majdi Jribi^(✉), and Faouzi Ghorbel^(✉)

CRISTAL Laboratory, GRIFT Research Group,
National School of Computer Science,
University of Manouba, 2010 La Manouba, Tunisia
amal.rihani@ensi-uma.tn,
{majdi.jribi,faouzi.ghorbel}@ensi.rnu.tn

Abstract. In this paper, we intend to propose a framework for the description and the matching of three dimensional faces. Our starting point is the representation of the 3D face by an invariant description under the $M(3)$ group of translations and rotations. This representation is materialized by the points of the arc-length reparametrization of all the level curves of the three polar representation. These points are indexed by their level curve number and their position in each level. With this type of description we need a step of registration to align 3D faces with different expressions. Therefore, we propose to use a robust version of the iterative closest point algorithm (ICP) adopted to 3D face recognition context. We test the accuracy of our approach on a part of the BU-3DFE database of 3D faces. The obtained results for many protocols of the identification scenario show the performance of such framework.

Keywords: 3D face · Description · Three polar representation · The arc-length reparametrization · Registration · Hausdorff · ICP · BU-3DFE

1 Introduction

The automatic 3D shape recognition has known a growing interest during the last years in the pattern recognition field. Recently, the 3D data become active especially with the 3D acquisition materials improvement and the big computer capacity in the term of calculations. Therefore, the quality and the resolution of 3D meshes become better. In addition, 3D data permit to overcome the problems often encountered in 2D data. In fact, 2D data need an invariance under the perspective transformations while the 3D data surfaces need only the invariance under the Euclidian transformations. But one of the major problems of 3D surfaces is the lack of a canonical parameterizations. This fact makes hard the matching procedure between 3D objects. In order to overcome as much as possible this limits many works propose to extract an invariant description from

3D surfaces under the initial parametrization. In the literature, the 3D shape description can be classified into two main categories: The global methods and the local ones.

Several 3D global surface descriptions were proposed in the literature. In this category, we can mention the cords histogram methods proposed by Paquet et al. [1]. Its consists on the extraction of the statistical characteristics from the cords of the 3D object. Osada et al. [2] proposed as a global description for the 3D surfaces, the 3D distribution forms method. This last one is obtained by a probability distribution of a 3D shape function.

For the second category of methods, a 3D local representation is extracted from a 3D objects. In this context, there are many local descriptors based on the curvature such as the Gaussian curvature proposed by Shw-wei et al. [3] which is used to describe the 3D faces. Also, Ganguly et al. [4] proposed to use a two pairwise of curvature analysis. The first pair is composed by the mean, and the maximum curvature and the second one corresponds to the minimum and the gaussian curvature. We can mention here, Bannour et al. [19] who presented a 3D surface description by a set of invariant points obtained from a set of uniform levels of the curvature values. Another kind of the local methods which based on the construction of the geodesic level curves around a feature point are used to represent the 3D surfaces. [5–7] proposed to describe the 3D surface by a set of geodesic level curves generated from a one reference point qualified by the unipolar representation. Other works proposed to use the representation based on many reference points in order to overcome the problem of the instability in the case of error of the reference point extraction. Ghorbel et al. [8] proposed to use the bipolar representation. It is obtained from two reference points. It consists on the levels of the superposition of the two geodesic potentials generated from two reference points. In this context, Jribi et al. [9] proposed to extend this representation to the three polar one based on the superposition of three geodesic potentials from three reference points instead of two.

The majority of these description methods require a registration step in order to estimate the variation between two shapes and to align them. In the literature, the registration methods between 3D shapes can be classified into two major categories. The first type is based on the local geometry to construct a valid hypotheses of mappings. In this category, we can classify the registration methods based on Hough transform and Hashage tables [15–18]. The second type performs the mapping by iterative algorithms. We can mention here the works of Bes et al. [12] who used an iterative techniques to extract the matched points. In this paper, we intend to propose a 3D face recognition technique based on two stages: The first one consists on the proposition of an invariant 3D face description. The second stage is a step of alignment of the 3D surface by a novel robust version of ICP [12].

Thus, this paper will structured as follows: we present in the second section a brief recall of the proposed representation. The implementation steps of the proposed representation on 3D faces are described in section three. The used similarity metric to compare between two shapes and the novel robust version

of ICP are detailed in the fourth section and finally, we test the accuracy of our representation for the identification scenario on a part of the BU-3DFE database of 3D faces in the last section.

2 Brief Recall of the Proposed Representation

In this paper, we propose to describe the 3D surfaces by an accurate, finite, and invariant set of points under the geometrical transformations of the $M(3)$ group of translations and rotations. This description is proposed by Rihani et al. [14]. It is obtained by two steps: (i) The first step consist on the construction of the three polar representation proposed by Jribi et al. [9]. (ii) in the second step a geometric arc-length reparametrization of each level of the three polar representation should be performed. We describe in the rest of the section the two steps cited above.

In the rest of the section, we consider that a 3D object as a 2D-differential manifold denoted by S .

2.1 The Construction of the Three Polar Representation

Let denote by U_r the function that computes for each point p of S the length of the geodesic curve joining it to the point r . The three polar representation consists on the superposition of three geodesic potential generated from three reference points. Therefore let denote by p_1, p_2, p_3 three reference points of S $U_{p_1}, U_{p_2}, U_{p_3}$ their corresponding geodesic potentials and U_s the sum of these three geodesic potentials. Thus, the three polar representation that we denote by $M^k(S)$ corresponds to the set of k level curves where each level curve C^{λ_i} is composed by a set of points having the sum of the three geodesic potential U_s equal to λ . It can be formulated as follows:

$$M^k(S) = \{C^{\lambda_i}\}_{i=1..k} \quad (1)$$

where

$$C^{\lambda_i} = \{p \in S, U_s(p) = \lambda_i\} \quad (2)$$

2.2 Geometric Arc-Length Reparametrization

After the construction of the three polar representation, the 3D surface S is presented by a collection of level curves $\{C^{\lambda_i}\}$. A curve parametrization $\{C^{\lambda_i}(t)\}$ is an 1-periodic function of a continuous parameter t defined by:

$$\begin{aligned} C^{\lambda_i} : [0, 1] &\rightarrow \mathbb{R}^3 \\ t &\mapsto [x(t), y(t), z(t)]^t \end{aligned} \quad (3)$$

It's well known that the same parametric curve C^{λ_i} can have many parameterizations. This due to parametrization dependance on the position, the orientation

of the used curve and the speed we go over. In order to overcome this problem, we propose to use a \mathbb{G} invariant reparametrization of each curve of the three polar representation where \mathbb{G} is a group of geometrical transformations applied to a curve.

In our context, \mathbb{G} corresponds to the $M(3)$ group formed by the \mathbb{R}^3 rotations and translations. This group of transformations preserves the length of the curve however the speed we go over the curve affects its parametrization. Therefore, we carry out an arc-length reparametrization of a 3D curve C^{λ_i} in order to cover it with the same speed. The arc-length reparametrization is defined as follows:

$$S(t) = 1/L \int_0^t \sqrt{x(t)^2 + y(t)^2 + z(t)^2} dt, t \in [0, T] \quad (4)$$

Here, L denotes the length of the level curve C^{λ_i} .

3 The Application of the Proposed Representation on 3D Faces Meshes

Since the 3D faces known a growing interest for the identities determination especially after the many terrorist acts occurred around the world, we implement this novel representation on this type of data. In practice, the 3D surface corresponds to a discrete mesh. We will start by the construction of the three polar representation on the 3D faces. As mentioned before the three polar representation is based on the three reference points. In our case, the out corner of the eyes and the noise tip are used as reference points. For the automatic extraction of the reference points, we use an approach based on a curvature analysis of 3D faces proposed by Szeptycki et al. [21]. Then, for each reference point we compute its geodesic potential. In the discrete case, the computation of a geodesic potential generated from a reference point corresponds to the computation of the geodesic curves between the reference point and the other points of the 3D face. Here, we use the fast marching algorithm [13] for the computation of the geodesic path between each pairs of points. The three polar representation is composed by a set of discrete level curves. Each level curve of value λ can be represented by a set of vertices. The sum of three geodesic potentials of each vertex should belongs to $[\lambda - \epsilon, \lambda + \epsilon]$ it can formulated as follows:

$$C^\lambda = \{P \in S, \lambda - \epsilon \leq U_3(P) \leq \lambda + \epsilon\} \quad (5)$$

where ϵ is a real positive value chosen according to the resolution of mesh to avoid the intersections between successive level curves.

After the construction of the geodesic level curves of the three polar representation, we perform the approximation of these curves by the B-spline function. Finally, we realize the arc-length reparametrization procedure for each level curve of the tree polar representation. The obtained points are equidistant and invariant under the $M(3)$ group of translations and rotations. Each point is defined

by its level number value and its position in that level. In fact, the 3D face can be defined for N levels of the three polar level curve by:

$$\hat{M}^N(S) = \{p_{ij}^S \in \mathbb{R}^3\}, i \in [1..N], j \in [1..L] \quad (6)$$

where N is number of the geodesic level curves of the three polar representation and L is number the points by level. In Fig. 1, we summarize all the steps of the proposed representation applied on the 3D faces.

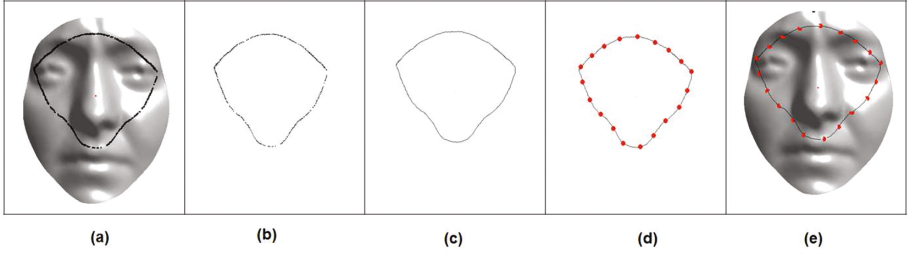


Fig. 1. The steps of the proposed approach applied to a 3D face. (a, b): The extraction of the three polar level curve. (c): Approximation of this level curve with the B-spline function. (d, e): The arc-length reparametrization of this level curve.

4 3D Faces Comparison

4.1 Hausdorff Shape Distance

In this work, we use the well known Hausdorff shape distance introduced by Ghorbel et al. [10, 11] for the recognition task between 3D shapes. All the possible parameterizations of surface are grouped on G . G can be \mathbb{R}^2 plane if the surface is open or \mathbb{S}^2 if it is closed. Let S_1 and S_2 be two 3D surface pieces diffeomorphic to G on which act the $M(3)$ group of geometrical transformations. The Hausdorff shape distance between S_1 and S_2 can be defined by:

$$\Delta(S_1, S_2) = \max(\rho(S_1, S_2), \rho(S_2, S_1)) \quad (7)$$

where :

$$\rho(S_1, S_2) = \sup_{g_1 \in M(3)} \inf_{g_2 \in M(3)} \|g_1 S_1 - g_2 S_2\|_{L^2}^2 \quad (8)$$

Since the $M(3)$ displacement group preserves this norm, the Hausdorff shape distance can be written as the following quantity:

$$\Delta(S_1, S_2) = \inf_{h \in M(3)} \|S_1 - h S_2\|_{L^2}^2 \quad (9)$$

The transformation between two shapes should be estimated in order to compute the correct value of the Hausdorff shape distance. We use in our context, a novel

robust version of the Iterative Closest Point algorithm to estimate the optimal transformation between faces. In this work, each 3D face is characterized by its 3D descriptor. Therefore a face is described by a set of infinite points obtained after the reparametrization of the tree-polar geodesic level curves.

4.2 Proposed Robust Version of ICP

In this work, we are interested on the problem of the 3D faces recognition. In this context, we generally need an elementary process of fine alignment which consists on the minimization of the global deviation between surfaces to compute the right distance value. But the major problem of such type of surfaces consists on the uncontrolled effects of the facial expressions. Therefore, we propose here, a robust version of the iterative closest point algorithm (ICP) adopted to this context. The ICP algorithm takes as input two 3D surfaces characterized by their points cloud. ICP is based on three main steps: (i) The first one consists on the matching procedure between the two sets of points. (ii) In the second step, the optimal rigid transformation is estimated. (iii) We apply finally the estimated transformation to one of the sets of points. The main contributions of the proposed version of ICP are essentially in its two first steps.

Here, a 3D face is represented by a set of discrete points corresponding to the proposed descriptor. The descriptor of a 3D face S_1 is formulated as mentioned above by:

$$\hat{M}^N(S_1) = \{p_{ij}^{S_1} \in \mathbb{R}^3\}, i \in [1..N], j \in [1..L] \quad (10)$$

where N is number of the three polar level curve of the three polar representation and L is number the points by level.

Let consider two surfaces S_1 and S_2 and their respectively corresponding descriptors $\hat{M}^N(S_1)$ and $\hat{M}^N(S_2)$ are defined by:

$$\begin{aligned} \hat{M}^N(S_1) &= \{p_{ij}^{S_1} \in \mathbb{R}^3\} \\ \hat{M}^N(S_2) &= \{p_{ij}^{S_2} \in \mathbb{R}^3\}, i \in [1..N], j \in [1..L] \end{aligned} \quad (11)$$

First Step: Pairwise Points Matching. Bes et al. [12] determined that the matching step assumed 95% of the ICP's time. This fact shows that the efficiency of the ICP depends on the corresponding step. In our approach, the 3D surface is presented by a set of discrete points. These points are indexed by their level number value and their position in this level. The first contribution of the proposed robust version of the ICP derive directly from the three polar representation. In fact, the matching procedure is automatically obtained since each point $p_{ij}^{S_1}$ is matched to the point $p_{ij}^{S_2}$ of the second face. One the other hand, a correct correspondence is conditioned by having a unique way to obtain the starting point on each level curve. We use, therefore, the plane passing through the noise tip and the first level of the three polar representation (which correspondence to a invariant point) to detect the starting point in each level curve.

The intersection between this plane and the 3D surface in each level curve of the three polar representation corresponds to the starting points of each three polar level curve.

Second Step: Transformation Estimation. The second step of ICP consists on the estimation of the rigid transformation between $\hat{M}^N(S_1)$ and $\hat{M}^N(S_2)$ that we denote by \hat{T} . ICP algorithm is an iterative procedure minimizing the Mean Square Error (MSE). In practice, the rigid transformation should find a solution to the least squares defined by:

$$\hat{T} = \underset{T}{\operatorname{argmin}} \sum_i \sum_j e_{ij}^2 \quad (12)$$

where e_{ij} is the distance between the point $p_{ij}^{S_1}$ of S_1 and its corresponding point $p_{ij}^{S_2}$ of S_2 .

$$e_{ij}^2 = \|p_{ij}^{S_2} - T(p_{ij}^{S_1})\|^2 \quad (13)$$

Our approach is implemented on the 3D faces with different facial expressions. Since the rigid matching process is sensitive to the 3D shape deformations, we should consider this variation shape. In the present work, we propose to automatically associate different weights to the different points representing the 3D surface. In fact, only the points that are less influenced by the facial expressions will participate in this estimation step. To distinguish these points, we suggest to study the variation V_{ij}^k of each point $p_{ij}^{S_k}$ of the surface S_k from its centroid noted by C_{S_k} in all the surfaces. This variation corresponds to the distances between $p_{ij}^{S_k}$ and C_{S_k} . It can be defined by:

$$V_{ij}^k = (d(p_{ij}^k, C_{S_k})), \quad (14)$$

The weight value $W_{ij}^{S_k}$ given for each point $p_{ij}^{S_k}$ should qualify the quality of matching. Indeed, The more static the point is, the greater its weight should. Therefore, the weight W_{ij} for two corresponding points $p_{ij}^{S_1}$ and $p_{ij}^{S_2}$ for the two surfaces S_1 and S_2 can be formulated by:

$$W_{ij} = \frac{V_{max} - (V_{ij}^{S_2} - V_{ij}^{S_1})}{V_{max}} \quad (15)$$

where V_{max} is presented by:

$$V_{max} = \max_k (\max_{ij} (V_{ij}^k)), k \in [1..H] i \in [1..N], j \in [1..L] \quad (16)$$

where H is the number of the used 3D surfaces.

This equation shows that when the variation between two correspondent points tends to reach V_{max} the weight W_{ij} of p_{ij} tends to zero.

Thus, the novel transformation estimation should find a solution to the least squares defined by:

$$\hat{T} = \underset{T}{\operatorname{argmin}} \sum_i \sum_j W_{ij}^2 e_{ij}^2 \quad (17)$$

Seen that T is a rigid transformation, it can be decomposed on rotation and translation. Therefore, it can be defined as follows

$$\hat{T} = \operatorname{argmin}_T \sum_i \sum_j W_{ij}^2 (\|p_{ij}^{S_2} - T(p_{ij}^{S_1})\|^2) \tag{18}$$

The translation between the two sets of points is defined by:

$$\hat{t} = C_{S_2} - RC_{S_1} \tag{19}$$

where C_{S_2} and C_{S_1} are respectively the centroid of $\hat{M}^N(S_1)$ and $\hat{M}^N(S_2)$.

Once the rotation R is determined the translation can be derived. Therefore, we need firstly to estimate the rotation R . We place each set of points on its centroid landmark: $p_{C_{ij}}^{S_1} = p_{ij}^{S_1} - C_{S_1}$ and $p_{C_{ij}}^{S_2} = p_{ij}^{S_2} - C_{S_1}$. The optimal rotation is rewritten as follows:

$$\hat{R} = \operatorname{argmin}_R \sum_i \sum_j W_{ij}^2 \|p_{C_{ij}}^{S_2} - R(p_{C_{ij}}^{S_1})\|^2 \tag{20}$$

5 Experiments and Discussion

Here, we perform experiments based on the novel version of ICP applied to the reparametrized level curves for the identification scenario. For the experimentation, we used a part of the BU-3DFE database [20]. This portion is composed by 700 faces corresponds to the first magnitude level of the six facial expressions and the neutral face of all the subjects of the database(100 subjects). We run

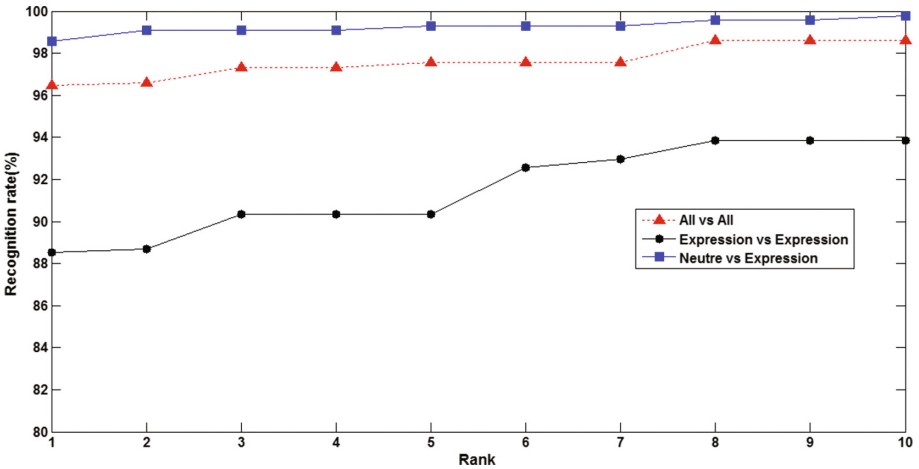


Fig. 2. The CMC curve of the proposed approach for the scenarios: All vs All, Expression vs Expression and Neuter vs Expression

the experiments with three protocols: (i) The first one is All vs All. It consists on the comparison of each face of the database to all the others. (ii) Expression vs Expression is the second protocol. This one corresponds to the comparison between each expression of the database and all the other expressions. (iii) Neuter vs Expression protocol is used to compare each 3D neutral face with the 3D faces with expression. Figure 2 shows the Cumulative Matching Curves of the proposed 3D representation under the three protocols cited above. The obtained rank-one recognition rates are about 96.48% for All vs All protocol, 88.53% for Expression vs Expression and 98.65% for Neuter vs Expression.

6 Conclusion

We introduced in this work a new approach for the recognition of the 3D faces. This approach consists on a novel robust version of the ICP algorithm. This proposed ICP is based on the three polar representation proposed in [9] and it is adopted to the variation of the facial expression on the 3D faces. The obtained rates for the three protocols of the identification scenario show the performance of the proposed framework.

We propose in the future work to experiment the proposed approach on the standard database of 3D faces FRGCv2. We intend also to compare the proposed ICP with ICP's variants.

References

1. Paquet E., Rioux M.: A query by content system for three-dimensional model and image databases management. In: The 17th conference on Image and Vision Computing, pp. 157–166 (1999)
2. Osada, R., Funkhouser, T., Chazelle, B., Dobkin, D.: Shape distributions. *ACM Trans. Graph.* **21**(4), 807–832 (2002)
3. Shinagawa, Y., Kunii, T.-L., Kergosien, Y.-L.: Surface coding based on morse theory. *IEEE Comput. Graph.* **11**, 66–78 (1991)
4. Ganguly, S., Bhattacharjee, D., Nasipuri, M.: 3D face recognition from range images based on curvature analysis. *ICTACT J. Image Video Process.* **4**(3), 748 (2014)
5. Samir, C., Srivastava, A., Daoudi, M.: Three dimensional face recognition using shapes of facial curves. *IEEE Trans. Pattern Anal. Mach. Intell.* **28**(11), 1858–1863 (2006)
6. Srivastava, A., Samir, C., Joshi, S.H., Daoudi, M.: Elastic shape models for face analysis using curvilinear coordinates. *J. Math. Imaging Vision* **33**(2), 253–265 (2008)
7. Gadacha, W., Ghorbel, F.: A new 3D surface registration approach depending on a suited resolution: application to 3D faces. In: *IEEE Mediterranean and Electrotechnical Conference (MELECON)*, Hammamet, Tunisia (2012)
8. Ghorbel, F., Jribi, M.: A robust invariant bipolar representation for R3 surfaces: applied to the face description: Springer. *Ann. Telecommun.* **68**(3–4), 219–230 (2013)

9. Jribi, M., Ghorbel, F.: A stable and invariant three-polar surface representation: application to 3D face description. In: WSCG 2014, the 22nd International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision, Republic (2014)
10. Ghorbel, F.: A unitary formulation for invariant image description: application to image coding **53**(5–6), 242–260 (1998). Special issue *Annales des telecommunications*
11. Ghorbel, F.: Invariants for shapes and movement. Eleven cases from 1D to 4D and from Euclidean to Projectives (French version), Arts-pi edn., Tunisia (2012)
12. Besl, P.J., McKay, N.D.: A method for registration of 3-D shapes. *IEEE Trans. Pattern Anal. Mach. Intell.* **14**(2), 239–256 (1992)
13. Sethian, J.A.: A fast marching level set method for monotonically advancing fronts. *Proc. Nat. Acad. Sci.* **93**, 1591–1595 (1996)
14. Rihani, A., Jribi, M., Ghorbel, F.: A novel accurate 3D surfaces description using the arc-length reparametrized level curves of the three-polar representation. In: WSCG 2016, the 24th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision, Republic (2016)
15. Ayache, N.: Computer vision applied to 3D medical imagery: results, trends and future challenges. In: *Proceedings of the 6th Symposium on Robotics Research*. MIT Press, also Inria Tech. (1993)
16. Faugeras, O., Hebert, M.: The representation, recognition and positioning of 3d shapes from range data. In: *Proceedings of the 8th International Conference On Artificial Intelligence*, Karlsruhe, BRD, pp. 996–1002, August 1983
17. Rigoutsos, I., Hummel, R.: Robust similarity invariant matching in the presence of noise: a data parallel approach. In: *Proceedings of the 8th Israeli Conference on Artificial Intelligence and Computer Vision*
18. Gueziec, A., Ayache, N.: Smoothing and matching of 3D-space curves. In: *Proceedings of the Second European Conference on Computer Vision Santa Maragherita Ligure, Italy, May 1992*
19. Bannour, M.T., Ghorbel, F.: Isotropie de la représentation des surfaces; Application à la description et la visualisation d'objets 3D. In: *RFIA 2000*, pp. 275–282 (2000)
20. Lijun, Y., Xiaozhou, W., Yi, S., Jun, W., Matthew, J.: A 3D facial expression database for facial behavior research. In: *The 7th International Conference on Automatic Face and Gesture Recognition*, pp. 211–216 (2006)
21. Szeptycki, P., Ardabilian, M., Chen, L.: A coarse-to-fine curvature analysis-based rotation invariant 3D face landmarking, In: *The IEEE 3rd International Conference on Biometrics: Theory, Applications, and Systems, BTAS 2009* (2009)