TEXTURE-BASED FUZZY SYSTEM FOR ROTATION-INVARIANT CLASSIFICATION OF AERIAL ORTHOIMAGE REGIONS

M. A. Barrera*, M. T. C. Andrade, Hae Yong Kim

Escola Politécnica, Universidade de São Paulo, Brazil - (mauriciob, mtcandrade)@usp.br, hae@lps.usp.br

KEY WORDS: Aerial image, Texture classification, Fuzzy Logic, Land Use, Orthoimage.

ABSTRACT:

Orthoimages are aerial images where feature displacements and scale variations have been removed. This type of images is widely used to calculate areas, determine land cover and land use, among others. This paper introduces a rotation-invariant classification model for three common orthoimage regions: city, sea and forest areas, using only texture information (without color information). Our classification model analyzes small sub-images (for example, of 20x20 pixels) to determine their region classes. Our model is based on a Fuzzy Inference System (FIS) constructed over a set of new rotation-invariant texture features. The features are extracted using two rotation-invariant versions of the well-known grayscale co-occurrence matrix (GLCM). Rotation-invariance is a desirable property of orthoimage classification systems, because the aerial images can be taken from different angles. We executed tests on samples from the three regions, including several rotated versions. These experiments show that our system reaches 100% of correct classification rate for our image test database. This correct classification rate is far superior to the rate obtained using the classical GLCM without the rotation-invariant property. Our classifier is robust to images that contain small areas that do not belong to the overall region type. The results demonstrate that our model offers a reliable rotation-invariant orthoimage region classification.

1. INTRODUCTION

Orthorectified digital aerial photographs (orthoimages) are aerial images where feature displacements and scale variations have been removed. In this paper, we demonstrate experimentally that it is possible to classify forest, city and sea orthoimage regions using only texture information, without color or spectral information. We present a Fuzzy Inference System (FIS) for orthoimage region classification. We propose and use new rotation-invariant texture features. These features are modified versions of the well-known texture descriptors extracted from grayscale co-occurrence matrix (GLCM) (Haralick et al., 1973).

There are some papers that use FIS for satellite images classification, for example (Binaghui et al., 1997; Wei, 2010). Also, there are some rotation-invariant texture descriptors in the literature, for example wavelet-based methods (Pun et al., 2003) and local binary pattern methods (Song et al., 2010). Machine learning techniques have also been used to achieve partial rotation invariance and these approaches have been applied to the traditional GLCM (Salem at al., 2011). One of the authors participated in the proposal of a technique that makes GLCM intrinsically rotation-invariant (Ito et al., 2009). We will use this technique to obtain rotation-invariant orthoimage classification.

2. ROTATION-INVARIANT CO-OCCURRENCE MATRIX

The grayscale co-occurrence matrix (GLCM) is widely used as image texture descriptor. We modify the GLCM to make it rotation-invariant, obtaining two new co-occurrence matrices: radial and circular. We extract the rotation-invariant texture descriptors from the new matrices. Baraldi et al. examined the GLCM texture descriptors defined in the literature and concluded that some of them were more statistically significant. These descriptors are the homogeneity, contrast, entropy, correlation, energy and variance (Baraldi et al., 1995). We make use of the first three descriptors.

The elements of a co-occurrence matrix provide information on the frequency of gray-level transitions between neighboring pixels in an image. Mathematically, a $L_g \times L_g$ co-occurrence matrix $P_{(\Delta x,\Delta y)}$ defined over a $n \times m$ grayscale image I with L_g gray-levels and parameterized by an offset $(\Delta x, \Delta y)$ is:

$$P_{(\Delta x,\Delta y)}(i,j) = \sum_{p=1}^{n}\sum_{q=1}^{m}\begin{cases}1, \text{ if } I(p,q)=i \text{ and } I(p+\Delta x, q+\Delta y)=j \\ 0, \text{ otherwise}\end{cases} \quad (1)$$

Clearly, the original co-occurrence matrix is not rotation-invariant. In order to make it rotation-invariant, we define below the circular and radial co-occurrence matrices.

Circular co-occurrence matrix: Let us denote $F_r(p,q)$ as the mean gray-level on the circle ring centered at (p, q) with radius r (Figure 1). The circular co-occurrence matrix $P_{(r_i,r_e)}$ with internal radius r_i and external radius r_e is:

This work was financially supported by CAPES.
* Corresponding author.

$$P_{(r_i,r_e)}(i,j) = \sum_{p=1}^{n}\sum_{q=1}^{m}\begin{cases}1, \text{ if } F_{r_i}(p,q)=i \text{ and } F_{r_e}(p,q)=j \\ 0, \text{ otherwise}\end{cases} \quad (2)$$

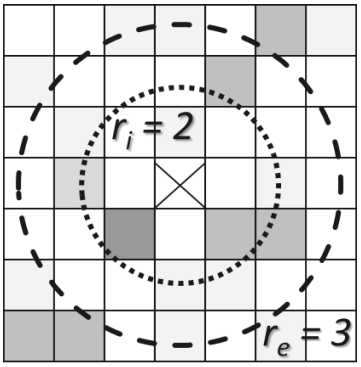


Figure 1. Calculation of average grayscales on circular rings, with radii $r_i = 2$ and $r_e = 3$.

Radial co-occurrence matrix: Let us denote $F_{(r,\theta)}(p,q)$ as the mean gray-level on radial line centered at (p, q) with radius r and inclination θ (Figure 2). The radial co-occurrence matrix with radius r and N uniformly-spaced radial lines is:

$$P_{(r,N)}(i,j) = \sum_{p=1}^{n}\sum_{q=1}^{m}\sum_{l=0}^{N-1}\begin{cases}1, \text{ if } F_{\left(r,\frac{2\pi l}{N}\right)}(p,q)=i \text{ and } F_{\left(r,\frac{2\pi(l+1)}{N}\right)}(p,q)=j \\ 0, \text{ otherwise}\end{cases} \quad (3)$$

The sum (l+1) must be computed modulus N. In this work, we used r = 5 and N = 8. In order to enhance the robustness of the method, it is possible to use the weighted mean of some number of neighboring radial lines, instead of using the mean of only one radial line. This is roughly equivalent to low-pass filtering the image. Figure 2 depicts the use of the mean of four radial lines (external two black lines and central two gray lines) to compute each $F_{(r,\theta)}(p,q)$.

Once we calculate both co-occurrence matrices, they are normalized by dividing each element by the sum of the matrix elements:

$$p(i,j) = \frac{P(i,j)}{\sum_{i=1}^{L_g}\sum_{j=1}^{L_g} P(i,j)} \quad (4)$$

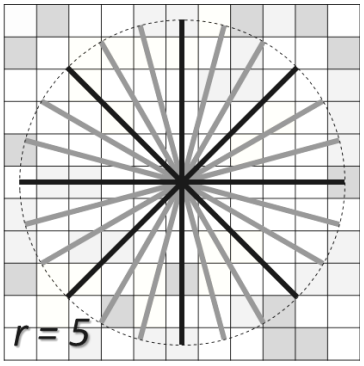


Figure 2. Calculation of average grayscales on N=8 directions with radius r=5. We used 4 radial lines to compute the average in each direction.

Then, we proceed to calculate the homogeneity, contrast and entropy descriptors (one for each matrix). The formulae for these descriptors are:

$$f_{hom} = \sum_{i=1}^{L_g}\sum_{j=1}^{L_g}\frac{1}{1+(1-j)^2}\,p(i,j) \quad (5)$$

$$f_{con} = \sum_{n=0}^{L_g-1} n^2 \left\{\sum_{i=1}^{L_g}\sum_{j=1}^{L_g} p(i,j)\right\}_{|i-j|=n} \quad (6)$$

$$f_{ent} = -\sum_{i=1}^{L_g}\sum_{j=1}^{L_g} p(i,j)\log\left(p(i,j)\right) \quad (7)$$

where $p(i,j)$ is the normalized co-occurrence matrix; $p_x(i) = \sum_{j=1}^{L_g} p(i,j)$ and $p_y(j) = \sum_{i=1}^{L_g} p(i,j)$ are the marginal distributions; μ_x and μ_y are the mean value of the marginal distributions; σ_x and σ_y are their standard deviations.

We obtain 6 rotation-invariant texture descriptors, three related to the circular co-occurrence matrix and three related to the radial version. We take each pair of descriptors (e.g.: circular and radial contrast) and derive three new features using the root of mean square of each pair:

$$f_i^{rms} = \sqrt{\frac{\left(f_i^{cir}\right)^2 + \left(f_i^{rad}\right)^2}{2}} \quad (8)$$

where $i \in \{hom, con, ent\}$; f_i^{cir} denotes the three circular descriptors and f_i^{rad} denotes the three radial descriptors. We use the three texture features f_i^{rms} in the remaining of this paper.

3. FUZZY INFERENCE SYSTEM

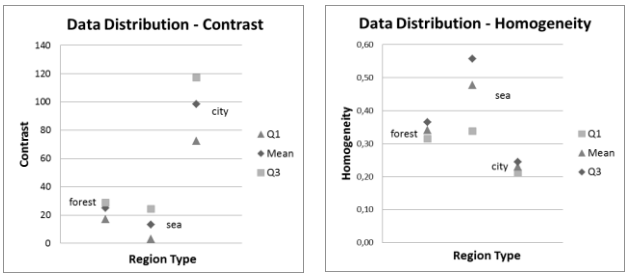
We acquired a set of orthoimages containing samples of the three region types: city, sea and forest for the training and tests. We

obtained grayscale images from the color ones by adding the red, green and blue components with weights 0.3, 0.59 and 0.11 respectively.

We gathered 56 images with an average size of approximately 360x240 pixels for training. We divided these training orthoimages in 20x20 pixel blocks, obtaining over 12000 training blocks. We also gathered another orthoimage database with 90 images (with approximately 400x400 pixels) for the tests. In order to validate the desired rotation-invariance characteristic, we generated 6 rotated versions of each test image, obtaining a total of 630 images. We used bilinear interpolation to rotate the images in angles $360°n/7$ where $n \in \{1, 2, ..., 6\}$.

The three *rms* features described in Subsection 3.1 were calculated for each block of the training images with 20x20 pixels. We used radii of 2 and 4 pixels in the circular co-occurrence matrix. We used a radius of 5 pixels and a set of angles varying in steps of 45° for the radial co-occurrence matrix. We plotted the mean, first and third quartiles of the data point distributions for each descriptor.

Figure 3 shows that the distribution of the contrast and homogeneity features falls into several well-defined intervals. The data distribution for the other two features (not shown) also follows the same behavior. By analyzing the distributions, we associated fuzzy intervals to each region. For example, the data points for a forest region would fall into a "low" contrast level and a "medium" homogeneity level.



(a) Contrast feature. (b) Homogeneity feature.

Figure 3. Data distribution of rotation-invariant features.

From this analysis, we associated the texture features to input linguistic variables and assigned linguistic labels to the data intervals. The possibility of belonging to one region type was associated to the range [0,1] in output linguistic variables. We derived trapezoidal membership functions and associated them to intervals of the rotation-invariant features. The universe of discourse for each input variable was set to the maximum possible range of values each feature can take. For the output variables, the universe of discourse was set to [0,100].

Each input linguistic variable can be used to distinguish three or two regions. For example, the homogeneity feature in Figure 3(b) is used to distinguish three classes: sea (high), forest (medium) and city (low). The contrast feature in Figure 3(a) is used to distinguish only two classes: city (high) and forest/sea (low). We denote $\mu(x)$ as the membership function associated to any input linguistic label; x belongs to the universe of discourse of the corresponding input linguistic variable and ρ is a vector that contains all the training data values associated to the linguistic label. We propose the equation below to model the membership function of any input linguistic label:

$$\mu(x) = \begin{cases} (x - \mu + 2\sigma)/(Q_1 - \mu + 2\sigma), & \mu - 2\sigma \le x < Q_1 \\ 1, & Q_1 \le x \le Q_3 \\ (\mu + 2\sigma - x)/(\mu + 2\sigma - Q_3), & Q_3 < x \le \mu + 2\sigma \\ 0, & \text{otherwise} \end{cases} \quad (9)$$

where σ is the standard deviation of ρ, μ is the arithmetic mean of ρ, Q_1 is the first quartile of ρ and Q_3 is the third quartile of ρ. This equation generates trapezoidal membership functions.

The membership functions of the output variables are three triangular fuzzy numbers associated, each of them, to the linguistic labels: not likely, likely and very likely. Figure 5 illustrates the input and output linguistic variables and the membership functions for each fuzzy set.

Finally, we created 15 fuzzy rules that characterize the relations between the feature value intervals and the associated outputs. We applied these rules under Mamdani's inference method (Lee, 1990) and used the centroid defuzzification method. Figure 4 shows three example rules that exploit the relations between the homogeneity and contrast descriptors. Compare Figures 3 and 4 to intuitively verify these relations. The complete set of rules is not presented due to space restrictions.

1. **IF** contrast is low **AND** homogeneity is medium **THEN** forest is likely
2. **IF** contrast is high **AND** homogeneity is low **THEN** city is likely
3. **IF** contrast is low **AND** homogeneity is high **THEN** sea is likely

Figure 4. Example of rules derived from the homogeneity and contrast features.

4. GENERAL STRUCTURE OF THE CLASSIFIER

The general scheme of the classifier is summarized in Figure 6. The training flow is illustrated with dashed arrows and the application flow is illustrated with solid arrows.

The training section initiates with the extraction of 20x20 pixel blocks from the training images. Then, we calculate the texture features for each training block and build a fuzzy inference system as presented in Subsection 3.2.

The application section also initiates with the extraction of 20x20 pixel blocks from the test image. We calculate the three texture features for each extracted block. The vector with these three features is used as the input of the fuzzy classifier. The region type that yields the highest possibility is associated to each block. If two or more regions yield the same and highest possibility, the block is associated to a null class. The final classification of the test image follows a simple rule: the region

type with the highest number of occurrences is elected as the land type of the test image.

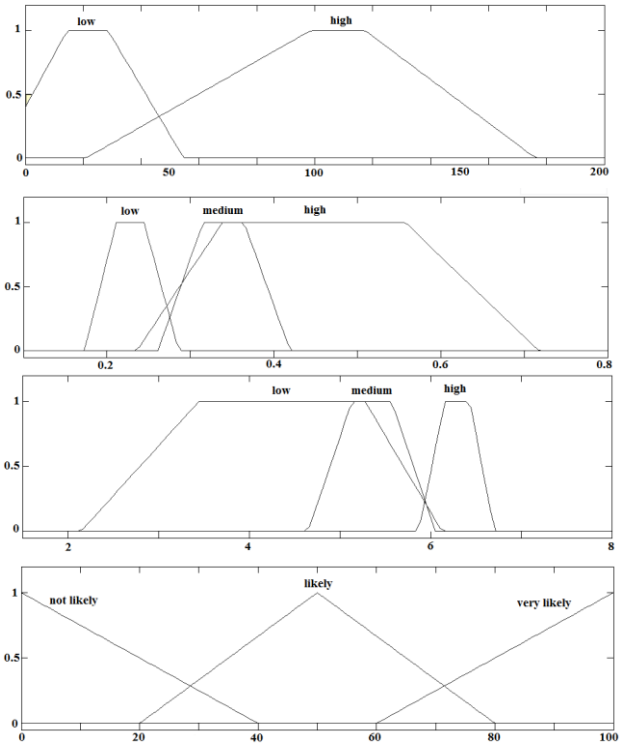


Figure 5. Input and output linguistic variables. From top to bottom: contrast, homogeneity, entropy and general output (city, forest, sea).

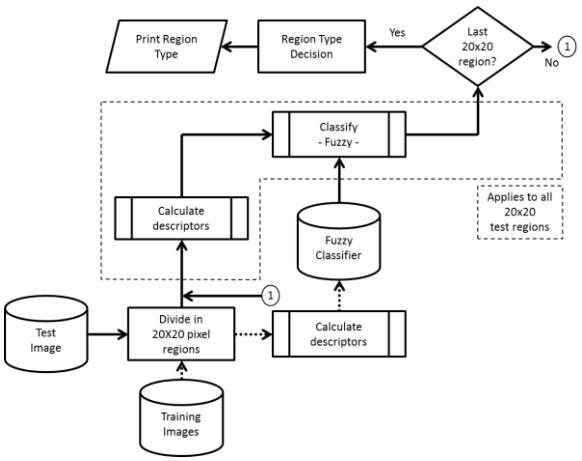
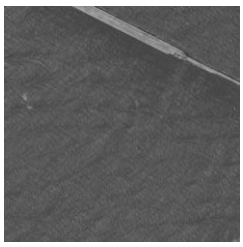


Figure 6. Texture analysis scheme for rotation-invariant region classification

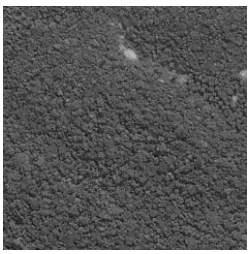
5. EXPERIMENTAL RESULTS

We implemented the described system in Matlab. We gathered 90 images representing city, forest and sea for the tests; this means 30 images for each region type with average size of 400x400 pixels. Figure 7 shows three examples. Two of them contain small secondary regions, as seen in Figures 7(b) and 7(c).

(a) City Region.



(b) Sea Region – Docks are visible.



(c) Forest Region – Small dirt patches are visible.

Figure 7. Examples of Non-Rotated Test Orthoimages.

From these 90 initial images, we generated 6 rotated versions for each (as stated at the beginning of Section 3), obtaining 630 test images. The rotated images have an average size of 260x260 pixels; we extracted the central area of the rotated images in order to eliminate the diagonal borders created by the rotations. For all the tests, we quantized the images from 256 to 128 grayscale levels.

We divide the results in two subsections. The first is concerned in classifying the complete images and the second deals with classifying each 20x20 pixel squares.

5.1 Classification of Complete Images

We first classified the 90 non-rotated orthoimages. We obtained no misclassifications, that is, the correct classification rate was 100%. We also tested the scheme for the remaining 540 rotated orthoimages. Again, we obtained no misclassifications. The total correct classification rate considering all 630 images was 100%.

To validate the effect of our rotation-invariant descriptors, we repeated the experiments using the classic horizontal grayscale co-occurrence matrix. We used our 630 image database to test the performance. We obtained a correct classification rate of only 25.25%.

5.2 Classification of Individual Region Blocks

In this subsection, we consider the classification of each 20x20 pixel block, instead of analyzing the classification of whole images.

Using all the test images, we classified around 127000 blocks. We calculated the classification rate for the blocks associated to the 90 non-rotated images and the 540 rotated images.

We obtained a correct block classification rate of 85.23% for the non-rotated images and a correct block classification rate of 85.24% for the rotated images. The total correct classification rate for all the test blocks was of 85.24%.

We illustrate below the classification results of individual blocks using some of our test images. To make the results visible, we superimpose the masks depicted in Figure 8 over the blocks.

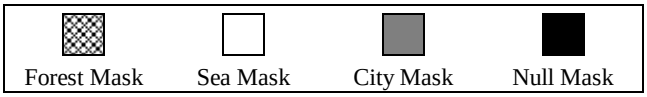
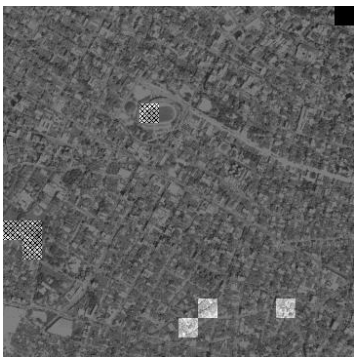


Figure 8. Block Classification Masks

Figure 9(b) shows the block classification for a city-type test image. There was one null classification and six misclassifications. Note that the classifier detected a forest-region on top of the visible stadium.

(a) Original Image

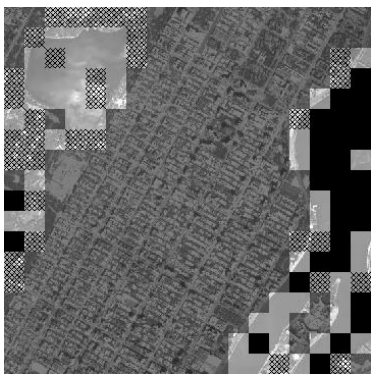


(b) Classified Image

Figure 9. Block Classification of a City Region

Figure 10(a) shows an image of a city region. There is a park and a lake to the left of the image; a river and an island stand to the right. According to Figure 10(b), the main city region was correctly classified (mid-gray area). The lake was classified as sea (clear-gray area). Some areas of the park were classified as forest (grid area). The two city-like spots to the bottom right were classified as such (mid-gray area). Finally, the river at right had some correctly classified areas and a considerable portion was not classified ("null mask").

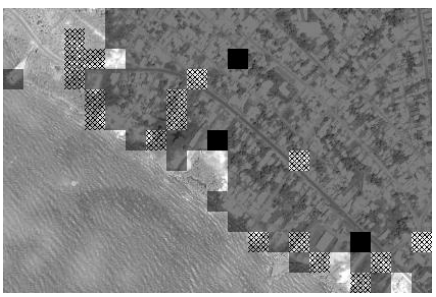
(a) Original Image



(b) Classified Image

Figure 10. Block Classification of a City Mixed Region

(a) Original Image



(b) Classified Image

Figure 11. Block Classification of a City and Sea Mixed Region

Figure 11 illustrates a coast city image with two regions (city and sea). In Figure 11(b), the majority of the city portion was correctly classified (mid-gray area), although there were some misclassifications. The sea region was correctly classified. The empty area in the upper left corner does not belong to any of the considered region types. The classifier associated it with sea and forest type regions.

6. DISCUSSION

The experimental results indicate that our rotation-invariant features describe robustly the texture information, even for rotated samples. Moreover, the inclusion of fuzzy logic provides a good platform for orthoimage classification.

The complete image correct classification percentages for the rotation-invariant approach (100%) and classical approach (25.25%) show clearly that our descriptors greatly enhance the performance for classifying rotated images.

The rotation-invariant features offer another advantage: it is possible to use a smaller training database. This is evidenced by the fact that our proposal coped much better with the unseen textures than the system based on the classic co-occurrence matrix. Our approach conserved the block correct classification percentages of the non-rotated and rotated images nearly identical (around 85.2%). In order to achieve a similar performance, the classic system would need extra training data with all possible rotation angles.

The block classification results show that it is possible to segment an orthoimage in the three regions using the proposed rotation-invariant classification.

We used images obtained from two sources: Google Maps and Microsoft Bing Maps. As a consequence, we built a system that correctly classifies orthoimages derived from aerial images captured by different devices. We achieved our results using only texture information derived from a single grayscale channel.

7. CONCLUSIONS

In this work, we proposed rotation-invariant texture features based on co-occurrence matrix and used them to classify orthoimages according to the land use. We proposed to use a fuzzy expert system. We tested our scheme in an image database and obtained a correct classification of 100%, while a similar system using classic co-occurrence matrix yielded only 25.25% of correct classification.

Our system proved to be robust to images containing small regions that do not belong to the predominant land type. The results demonstrate that our model offers a reliable rotation-invariant orthoimage region classification. This was achieved by using only grayscale texture information.

As future work, the rotation-invariant texture features could be calculated from the various bands of multispectral images.

REFERENCES

Baraldi, A., Parmiggiani, F., 1995. An Investigation of the Textural Characteristics Associated with Gray Level Co-occurrence Matrix Statistical Parameters. *IEEE T. Geosci. Remote*, 33(2), pp. 293-304.

Binaghi, E., Madella, P., Montesano, M., Rampini, A., 1997. Fuzzy Contextual Classification of Multisource Remote Sensing Images. *IEEE T. Geosci. Remote*, 35(2), pp. 326-340.

Haralick, R., Shanmugam, K., Dinstein, I., 1973. Textural Features for Image Classification. *IEEE T. Syst. Man Cyb.*, 3(6), pp. 610-621.

Ito, R., Kim, H., Salcedo, W., 2009. Classificação de Texturas Invariante a Rotação Usando Matriz de Co-ocorrência. In: 8th International Information and Telecommunication Technologies Symposium.

Lee, C., 1990. Fuzzy Logic in Control Systems: Fuzzy Logic Controller, Part II. *IEEE T. Syst. Man Cyb.*, 20(2), pp. 419-435.

Pun, C., Lee, M., 2003. Log-Polar Wavelet Energy Signatures for Rotation and Scale Invariant Texture Classification. *IEEE T. Pattern Anal.*, 25(5), pp. 590-603.

Salem, Y., Nasri, S., 2011. Rotation Invariant Texture Classification using Support Vector Machines. In: International Conference on Communications, Computing and Control Applications.

Song, C., Yang, F., Li, P., 2010. Rotation Invariant Texture Measured by Local Binary Pattern for Remote Sensing Image Classification. In: Second International Workshop on Education Technology and Computer Science.

Wei, W., 2010. Research on Remote Sensing Image Classification Based on Neuro-Fuzzy and Texture Analysis. In: Proceedings of the 29th Chinese Control Conference.