Weighted lightweight image retrieval method based on linear regression

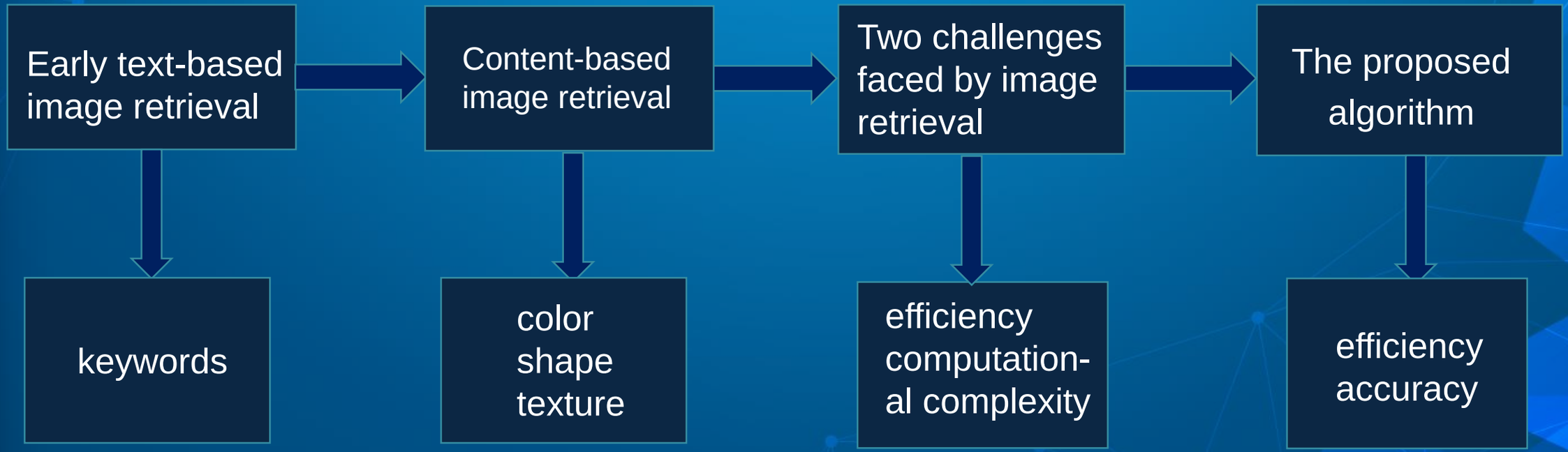
By Xiangqin Zheng
Xi'an University of Science and Technology

CONTENTS

1. Abstract

1. Research purpose and significance
2. Research method
3. Conclusion

2.Introduction



3. Preliminaries

RGB → HSV

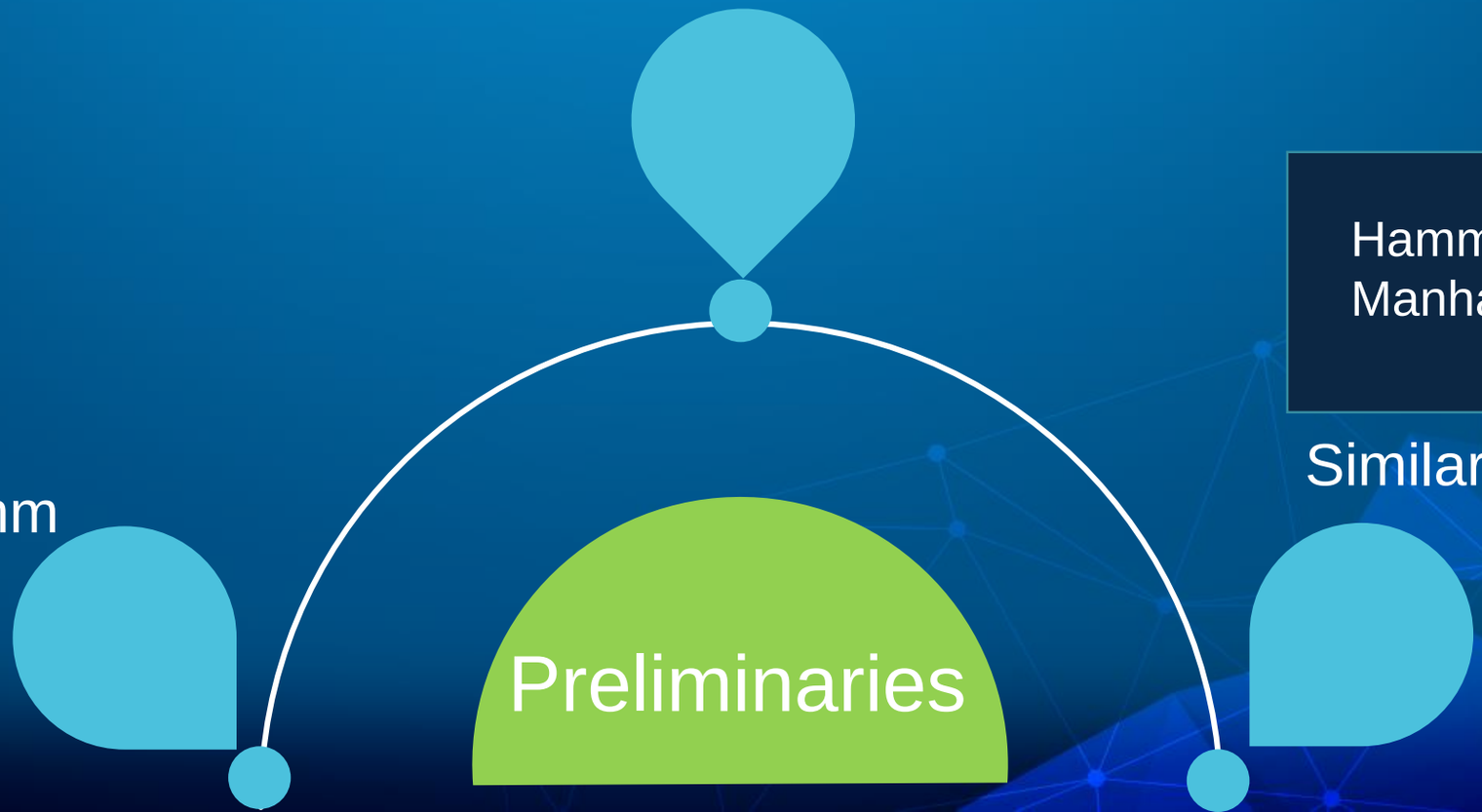
Color model conversion

Average hash
Perceptual hash
Difference hash

Image hash algorithm

Hamming distance
Manhattan distance

Similarity calculation



4.The proposed algorithm

1.
First, the image hash and color moments are merged with the weight:

$$z1 = \alpha_1 p + \beta_1 c$$
$$z2 = \alpha_2 a + \beta_2 c$$
$$z3 = \alpha_3 d + \beta_3 c$$

p, a, d are the Hamming distances

c is the Manhattan distance

$\alpha_1, \ldots, \alpha_3, \beta_1, \ldots, \beta_3$ are the weight

$z1, z2, z3$ are the new measures of similarity

2.

Then, linear regression analysis on the distance data obtained by color moments, perceptual hash and average hash is performed:

$$c' = \alpha'_0 + \beta'_0 p + \beta'_1 a \qquad (1)$$
$$s = \alpha' p + \beta' c'$$

Bring c' into s and merge similar items to obtain:

$$s = \beta' \alpha'_0 + \left(\alpha' + \beta' \beta'_0\right) p + \beta' \beta'_1 a \qquad (2)$$

where c' represents the new color moments represented by perceptual hash and average hash; $\alpha'_0, \beta'_0, \beta'_1$ represent the regression coefficients; α', β' are the new weights; s represents the result of fusing the perceptual hash with the new color moments and is used as a new similarity calculation method.

3.
Similarity acquisition

AH	PH	DH	CM
a_1	p_1	d_1	c_1
a_2	p_2	d_2	c_2
a_3	p_3	d_3	c_3
⋮	⋮	⋮	⋮
a_n	p_n	d_n	c_n

AH, PH, DH, and CM refer to average hash, perceptual hash, difference hash and color moments, respectively. $a_1,\ldots,a_n, p_1,\ldots, p_n, d_1,\ldots,d_n, c_1,\ldots,c_n$ respectively represent the distance or the similarity obtained by different algorithms in the retrieval process.

4. Multivariable linear regression model

A multivariable linear regression model can be expressed as：

$$Y = \gamma_0 + \gamma_1 X_1 + \gamma_2 X_2 + \ldots + \gamma_p X_p + \varepsilon \quad (3)$$

where γ and $X_i, i = 1, 2, \ldots, p$ represent the actual variables; γ_i, ε represent the regression coefficients and error terms, respectively. In this paper, Y represents the dependent variable CM and $X_i, i = 1, 2$ represents any two of the three independent variables AH, PH and DH. For example, we choose AH and PH for experiments.

The least squares estimate of the parameters is shown as below:

$$\hat{\gamma} = \left(X^T X \right)^{-1} X^T Y \quad (4)$$

In this paper, we use the least squares method for linear regression analysis to obtain the formula :

$$s = k0 + k1a + k2p$$

where a and p respectively represent the Hamming distance obtained by average hash and perceptual hash algorithm; k0, k1 and k2 are the regression coefficients obtained by linear regression; s is the new color feature expression.

5.Experimental results and analysis

1. Fusion of Hash and color

2. Fusion of hash and color moments obtained by linear regression

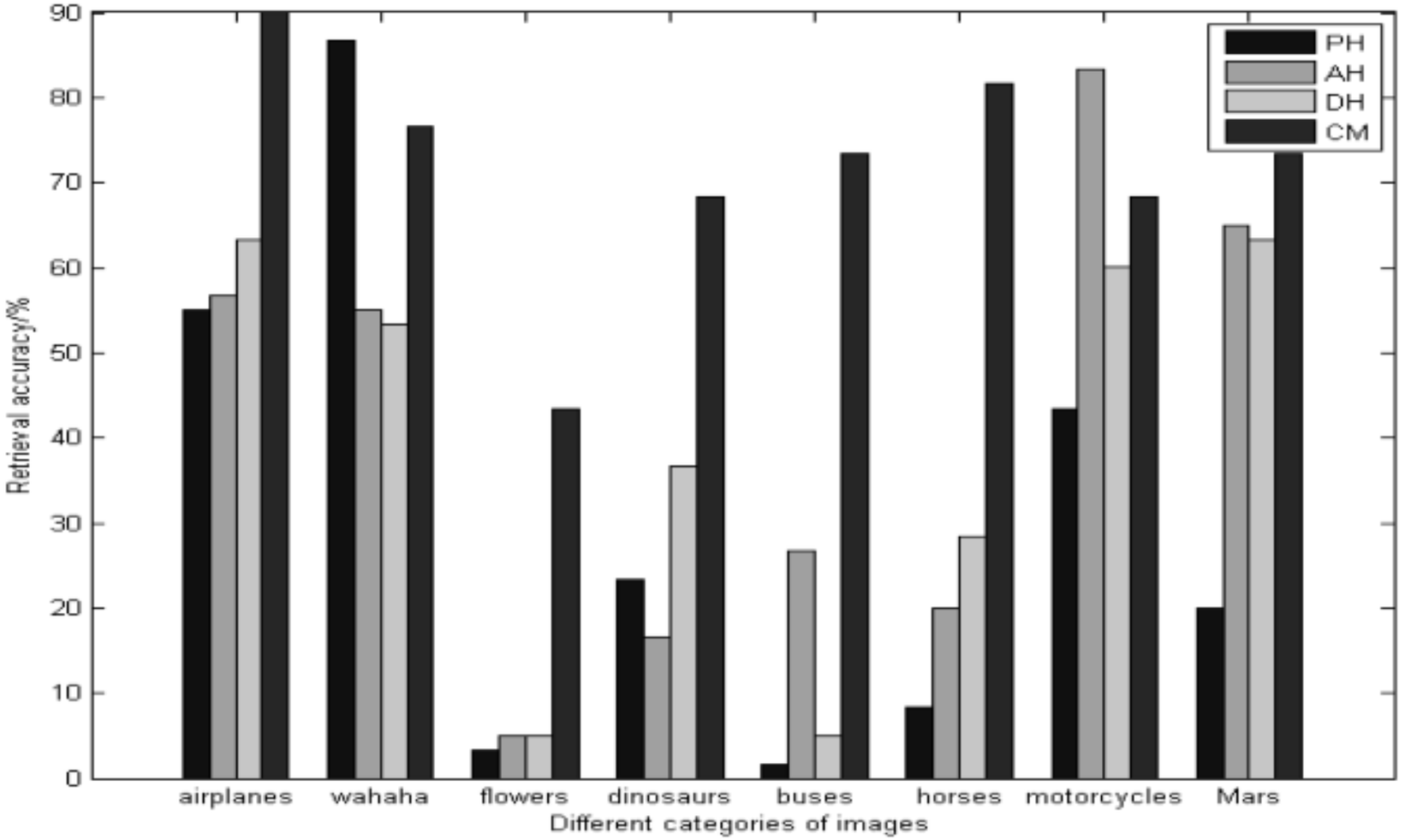
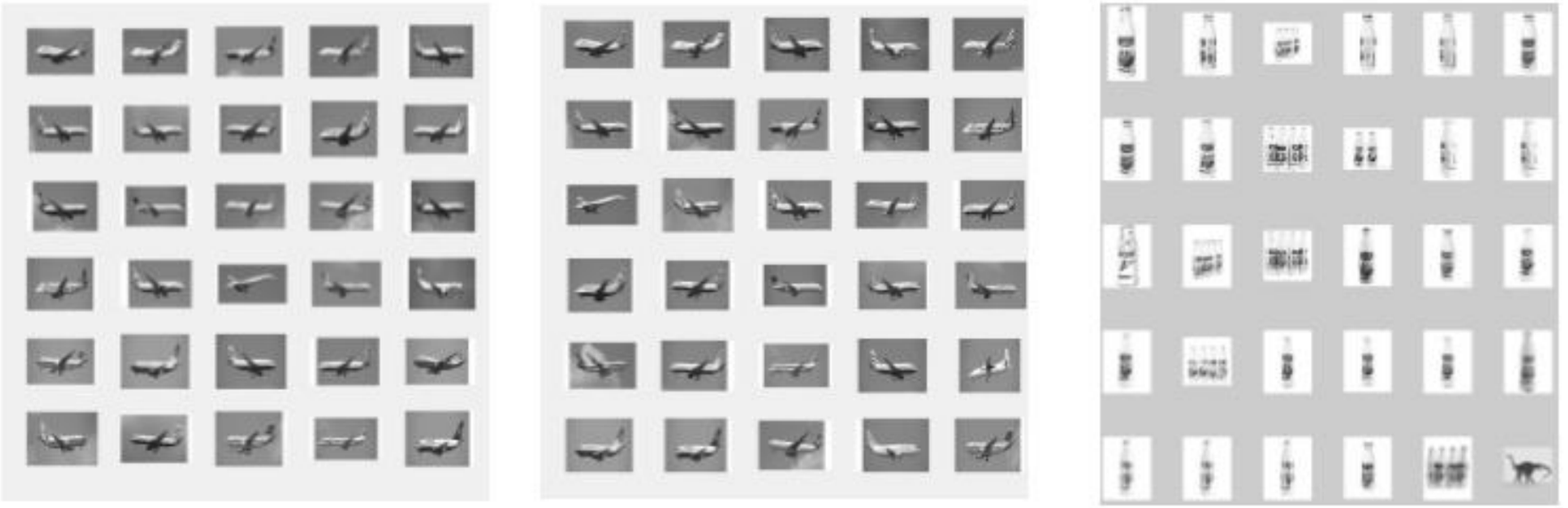


Fig. 1. The accuracy of different algorithms to retrieve different types of images



(a) (b) (c)

Fig. 2. Part of the results of retrieving the "airplanes" and "wahaha". (a) is part of the result of retrieving "airplanes" based on color moments, (b) is part of the retrieval result after the fusion of perceptual hash and color moments, and (c) is part of the search results after the fusion of perceptual hash and color moments obtained by linear regression

Comparison with related works

Method	Feature dimensions	Average retrieval time/s	The highest retrieval accuracy/%
IF	315	0.0861	80
TCHF	75	0.0263	< 60
TGMF	85	0.0510	< 80
LBP+DCT	469	0.1329	< 60
The proposed method	64	7.0	> 90

6.Conclusion

In this paper, a weighted lightweight image retrieval method based on linear regression is proposed. The image's perceptual hash, average hash and color moments are subjected to linear regression analysis. Then, the similarities obtained by faster hash algorithms are used to replace the color moments. Finally, the hash algorithm is merged with the new color moments to retrieve image. Experimental results show that the weighted lightweight image retrieval method based on linear regression has improved retrieval efficiency.

THANK YOU!