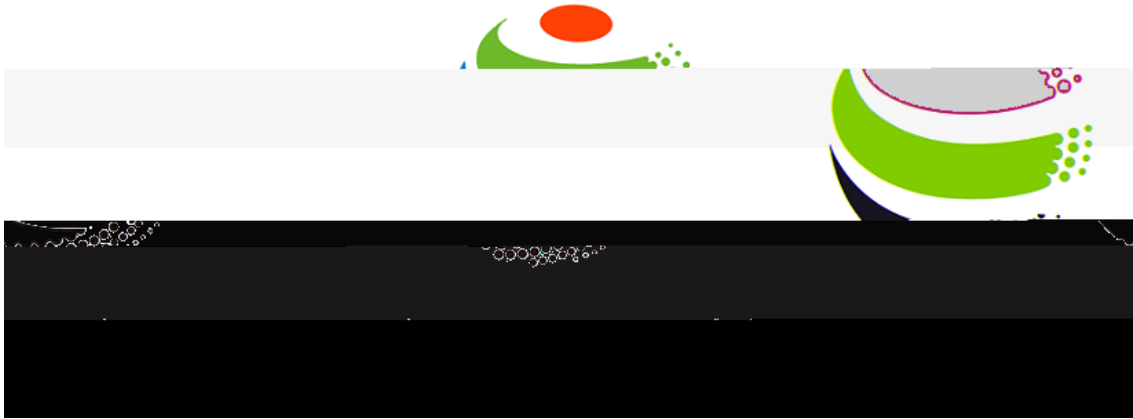


“

”



AVS

.....	1
.....	7
.....	11
.....	12
.....	13
.....	18
.....	27
.....	28

1	-	21
2	-	22
3	-	23
4	-	24
5	-	25
6		29

2015 8 21

“ ” , National Graduate Contest on Smart-city
Technology and Creative Design “
”

AVS

“ ”

“ ”

1.

2.

3.

4.

5.

6.

7.

8.

1.

2.

3.

1.

2.

3.

4.

5.

1

1.

2.

3.

4.

5.

www.smartcity-competition.com.cn

4

2

“

”

AVS

AVS

AVS

AVS

4

www.smartcity-competition.com.cn

http://page.renren.com/gra_smartcity



1.

2.

API

API

2

10

3.

4

CPU

CPU

GPU

U

2016	3	22
2016	4	1
2016	8	1
2016	8	20-22

1

1

2

3

1.

SDK

“ “ ”

1.

2.

CPU

GPU

3.

4.

5.

U

6.

7.

1

“

”

2

3

1

1 2

2

3

3 4

ppt flash

4

4

5

1

2

3

4

1

2

10-20

2016	3	22	-7	15					
2016	7	15		12:00					
2016	7	15	-8	1					
2016	8	1	-8	7					
2016	8	1	-8	20					
2016 8									

1

1-3

4-6

7-10

1

19

3

200

200

2000

1.

2.

2000

1.

2.

3.

2000

1.

2.

3.

4.

500

4

1

2

1

—

—

—

2

—

—

—

3

—

—

—

5

1

2

1

2

3

4

5

1

—

—

2

—
—
—
—
—
—
—
—
—
—

1

2

3

4

5

6

010-82337785

37

204

100191

010-62766552

5

501

pku_smartcity@126.com

010-52878507

PKU-SVD-B

PKU-SVD-B

“ ”

PKU-SVD-B

PKU-SVD-B

“ ”

PKU-SVD-B

PKU-SVD-B

PKU-SVD-B

1.

2.

PKU-SVD-B

PKU-SVD-B

3.

4.

5.

B Hongye Liu,

6.

_____ / _____

1.

“ ”

2.

3.

3

15

20

4.

5.

1.

avs

0 3
 1,3,5,7.....

2.

1.

3 3 3 8
 8

2.

3 3 3 8 12
 3
3 40 12

1.

avs

2.

SDK

evaluateType mediaFile VideoName

■ evaluateType

3

■ mediaGroup

2

MultiPedestrianTracking2

■ mediaCount

MultiPedestrianTracking1_peizhi.xml

8

2

mediaCount 2

■ VideoName

Items

■ startFrameNum

id

endFrameNum

endFrameNum

■ frameNum

■ l t r b id

id

■ type

1

2

<?xml version="1.0" encoding="gb2312" ?>

<Message Version="1.0">

<Info evaluateType="3" mediaGroup="MultiPedestrianTracking2" mediaCount="1" />

<TrackTarget>

<Items VideoName="yitiwest_8_1280x720_30" startFrameNum="57" endFrameNum="6299">

<Item id="9" type="2" l="333" t="101" r="356" b="154"/>

</Items>

<Items VideoName=" yitiwest_8_1280x720_30" startFrameNum="57" endFrameNum="1203">

<Item id="7" type="2" l="317" t="104" r="334" b="154" />

</Items>

<Items VideoName=" yitiwest_8_1280x720_30" startFrameNum="885" endFrameNum="6299">

```

    <Item id="3" type="2" l="373" t="106" r="391" b="154" />
  </Items>
  <Items VideoName=" yitiwest_8_1280x720_30" startFrameNum="1305" endFrameNum="2549">
    <Item id="2" type="2" l="310" t="101" r="328" b="143" />
  </Items>
</TrackTarget>
</Message>

```

3.

mediaFile	VideoName	SDK	evaluateType
■ evaluateType			
	3		
■ mediaCount		MultiPedestrianTracking1_	8
	3	mediaCount 3	
■ mediaFile		MultiPedestrianTracking1	
■ VideoName			Items
■ startFrameNum	id	endFrameNum	
	endFrameNum		
■ frameNum			
■ l t r b id			id
■ type	1	2	

```

<?xml version="1.0" encoding="gb2312" ?>
<Message Version="1.0">
  <Info evaluateType="3" mediaCount="2" meidiaFile="MultiPedestrianTracking1" />
  <Items VideoName="yitiwest_8_1280x720_30" startFrameNum="145" endFrameNum="6053">
    <Item frameNum="145">
      <Label type="2" l="621" t="147" r="653" b="224" id="9" />
    </Item>
    <Item frameNum="149">
      <Label type="2" l="618" t="147" r="651" b="224" id="9" />
    </Item>
    <Item frameNum="153">
      <Label type="2" l="617" t="148" r="649" b="225" id="9" />
    </Item>
  </Items>
</Message>

```

```

</Item>
<Item frameNum="157">
  <Label type="2" l="615" t="150" r="647" b="226" id="9" />
</Item>
</Items>
<Items VideoName="weiming_7_1280x720_30" startFrameNum="169" endFrameNum="6125">
  <Item frameNum="169">
    <Label type="2" l="1250" t="277" r="1279" b="438" id="9" />
  </Item>
  <Item frameNum="173">
    <Label type="2" l="1236" t="275" r="1279" b="437" id="9" />
  </Item>
  <Item frameNum="177">
    <Label type="2" l="1216" t="275" r="1279" b="441" id="9" />
  </Item>
  <Item frameNum="181">
    <Label type="2" l="1199" t="275" r="1258" b="449" id="9" />
  </Item>
  <Item frameNum="185">
    <Label type="2" l="1182" t="271" r="1248" b="443" id="9" />
  </Item>
</Items>
</Message>

```

1.

N

$T_n(n = 1, \dots, N)$

10

Ground

Truth

50%

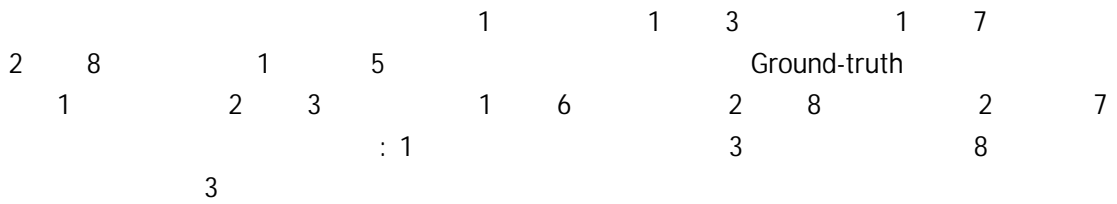
i

$$Recall(i) = \frac{\sum_{n=1}^N C_n^i}{\sum_{n=1}^N G_n^i},$$

$$Precision(i) = \frac{\sum_{n=1}^N C_n^i}{\sum_{n=1}^N D_n^i}$$

(1.1)

N $G_n^i (G_n^i < T_n)$ i Ground Truth
 n C_n^i i n
 D_n^i i n



$$Recall = 3/7, Precision = 3/5$$

$$\begin{aligned}
 Recall &= \frac{\sum_{i=1}^M \sum_{n=1}^N C_n^i}{\sum_{i=1}^M \sum_{n=1}^N G_n^i}, \\
 Precision &= \frac{\sum_{i=1}^M \sum_{n=1}^N C_n^i}{\sum_{i=1}^M \sum_{n=1}^N D_n^i}, \\
 Fscore &= \frac{2 * Recall * Precision}{Recall + Precision}
 \end{aligned} \tag{1.2}$$

M

2

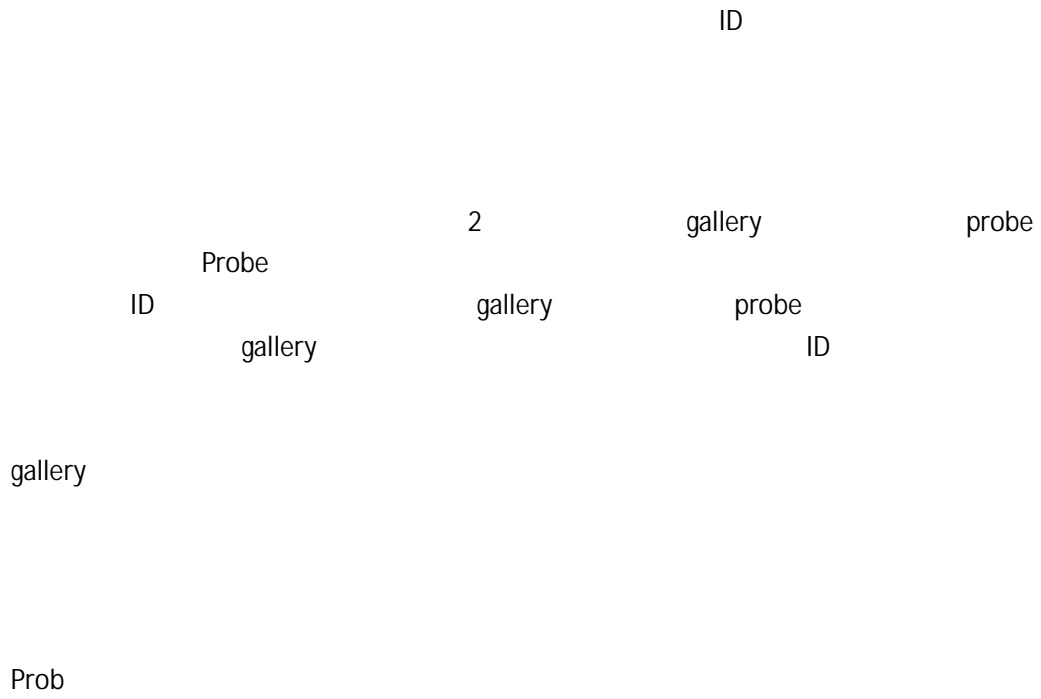
$$M = Fscore - L \tag{1.3}$$

L

$$\begin{aligned}
 &1.4 \\
 L &= \begin{cases} 0, & \text{if } t \leq 0; \\ t^2, & \text{if } t > 0. \end{cases} \tag{1.4}
 \end{aligned}$$

$$t = \frac{T - T_0}{120} \quad T \quad T_0$$

1.



2.



1.

	2	gallery	
groundtruth			1000

2.

		2	
1000	gallery		10
1100	gallery		10

1.

jpg

2.

	SDK	evaluateType
mediaFile VideoName		

■ evaluateType

4

■ mediaFile

dongnanmenwest_16_1920x1080_30

dongnanmenwest_16_1920x1080_30

■ imageName

■ score

0~1

■ l t r b id

id

<?xml version="1.0" encoding="gb2312"?>

<Message Version="1.0">

```

<Info evaluateType="4" mediaFile="dongnanmenwest_16_1920x1080_30" />
<Items>
  <Item imageName="13165.jpg">
    <Label id="30" score="0.800000" l="130" t="360" r="200" b="460" />
  </Item>
  <Item imageName="13167.jpg">
    <Label id="30" score="0.800000" l="150" t="370" r="230" b="460" />
  </Item>
  <Item imageName="13169.jpg">
    <Label id="30" score="0.800000" l="190" t="370" r="260" b="465" />
  </Item>
  <Item imageName="15161.jpg" />
</Items>
</Message>

```

1.

[0,1]

s ,

$$Recall(s) = \frac{t_p(s)}{t_p(s) + t_n(s)}, \tag{2.1}$$

$$Precision(s) = \frac{t_p(s)}{t_p(s) + f_p(s)}.$$

$t_p(s)$

$t_n(s)$

$f_p(s)$

1

2 2

ID ground-truth ID

Recall(s) *Precision(s)* ROC

Recall Precision

$$Precision = f(Recall); \tag{2.2}$$

C C

$$C = f(C). \tag{2.3}$$

1 2.2 2.3

10

2

2.3

2.3

ROC

2

$$M = C - L \tag{2.4}$$

L

$$L = \begin{cases} 0, & \text{if } t \leq 0; \\ t^2, & \text{if } t > 0. \end{cases} \tag{2.5}$$

$$t = \frac{T - T_0}{120} \quad T$$

T_0

2

1.

■

■

■

■

■

2.

70% K 30% M

■
■

1.

K 180

2.

60 120

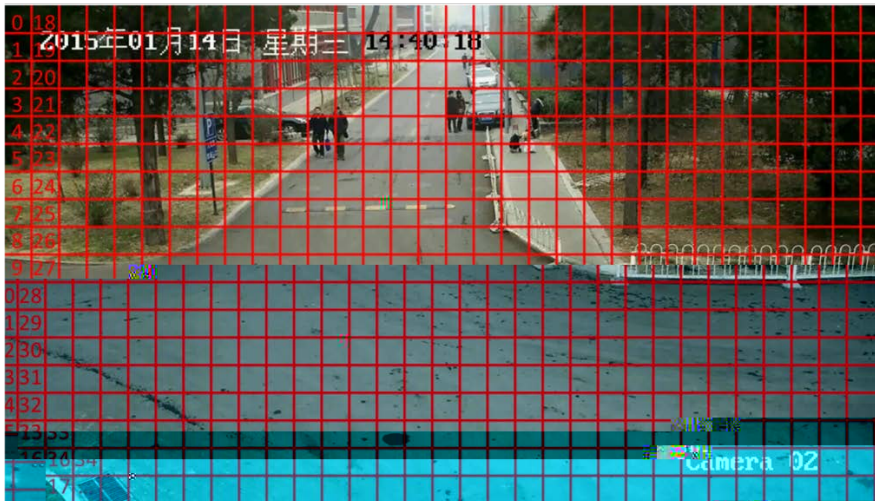
3.

R× C

0

R 18 C 32

18× 32 = 576



1.

avs

2.

SDK

evaluateType

mediaFile VideoName

■ evaluateType

5

■ mediaFile

2

AnomalyDetection2

- VideoName Items

- abn 1 0

- startFrame endFrame startFrame
endFrame 0 startFrame
endFrame -1

- frameNum

- blockId

```
<?xml version="1.0" encoding="gb2312" ?>
```

```
<Message Version="1.0">
```

```
  <Info evaluateType="5" mediaFile="AnomalyDetection2" />
```

```
  <Items VideoName="v0018_07_1920x1080_30.avs" abn="1" startFrame="34" endFrame="34">
```

```
    <Item frameNum="34">
```

```
      <Label blockId="124" />
```

```
      <Label blockId="78" />
```

```
      <Label blockId="258" />
```

```
    </Item>
```

```
  </Items>
```

```
</Items VideoName="v
```

<Items VideoName=" v0022_07_1920x1080_30.avs" abn="0" startFrame="-1" endFrame="-1" />
</Message>

1.

$$Precision_{clip} = \frac{\sum_i O_{clip}(i)}{\sum_i D_i} \quad (3.4)$$

$$D_i = i \cdot Precision_{clip}$$

F

$$F_{clip} = (1 + \beta) \frac{Recall_{clip} \times Precision_{clip}}{\beta Recall_{clip} + Precision_{clip}} \quad (3.5)$$

$$\beta = 0.75$$

$$Precision_{frame} = \frac{1}{N} \sum_{i=1}^N \frac{\sum_{j \in \{F_{i,p}, F_{i,d}\}} O_{Frame}(i,j)}{|\{F_{i,p}, F_{i,d}\} \cap \{G_{i,p}, G_{i,d}\}|} \quad (3.6)$$

N

Precision

$$Score = (1 + \beta) \frac{F_{clip} \times Precision_{frame}}{\beta F_{clip} + Precision_{frame}} \quad (3.7)$$

$$\beta = 0.5$$

2.

$$M = Score - L \quad (3.8)$$

L

$$L = \begin{cases} 0, & \text{if } t \leq 0; \\ t^2, & \text{if } t > 0. \end{cases} \quad (3.9)$$

$$t = \frac{T - T_0}{120} \quad T$$

T_0

2

1.

) query() ref N M N<=M ref(
query query ref
ID
id ref query
ref query

CVPR 2016

- *Hongye Liu, Yonghong Tian, Yaowei Wang , Lu Pang , Tiejun Huang. Deep Relative Distance Learning: Tell the Difference Between Similar Vehicles, In Computer Vision and Pattern Recognition(CVPR), 2016*

2.

ref
Mean Average Precision(MAP)@K(K=200)



8.44
ID

1.

	38998	5043		ID
0		250	7	

2.

16742	2000	,	ID	ref	query
	query		ref		

1.

jpg

2.

	SDK	evaluateType
mediaFile	VideoName	
■ evaluateType		
	6	
■ mediaFile		
	vehicle_retrieval_val	
■ Item	imageName	
■ Item	gallery	ID
(K=200	".jpg")

```

<?xml version="1.0" encoding="gb2312"?>
<Message Version="1.0">
  <Info evaluateType="6" mediaFile="vehicle_retrieval_val" />
  <Items>
    <Item imageName="012321 ">
      0292851 0110741 0173591 0092564 0286241 0192567 0340982 ...
    </Item>
    <Item imageName="003467 ">
      0387241 0023986 0283751 0230114 9806431 8823012 2389102 ...
    </Item>
    <Item imageName="13169 ">
      3727192 0387654 0007942 0009866 0120397 0485764 1200341 ...
    </Item>
    .....
  </Items>
</Message>

```

1.

MAP@K(K=200) P, AP@K, MAP@K

Precision query ref N

n $P(n) = \frac{n}{N}$

Average Precision AP@K query k $AP_k =$

$\sum_{1 \leq i \leq N} P_i / \min\{M, N\}$ M ref

Mean Average Precision(MAP@K) K query Average Precision

$MAP = \sum_{1 \leq k \leq K} AP_k / K$

1 probe Precision 0,

2 gallery 4-6

probe gallery

□ :

■ MAP@10

■ k query ID 1 ref 5 ID 1

■ 10 ID : 1, 3, 1, 2, 9, 8, 1, 3, 2, 4

■ $P(4) = 2/4 = 0.5$

■ $AP_k@10 = (1/1 + 2/3 + 3/7) / \min\{10, 5\} = 0.419$

■ $MAP@10 = \sum AP_k / K$

2

L

$$L_t = 0, \quad \text{if } t \leq$$

(3.9)

1.

2.

1.

2.

10

1.

yuv

avs

2.

AVS2
AI, LDB, LDP, RA

1.

BD-rate/BD-PSNR

Qp 22,27,32,37,42

BD-rate BD-PSNR BD-rate

BD-PSNR

PSNR BD-PSNR Qp

$$PSNR = 10 \times \log_{10} \left(\frac{(2^n - 1)^2}{MSE} \right)$$

MSE 4 rate-PSNR $(rate_{Qp}, PSNR_{Qp})$

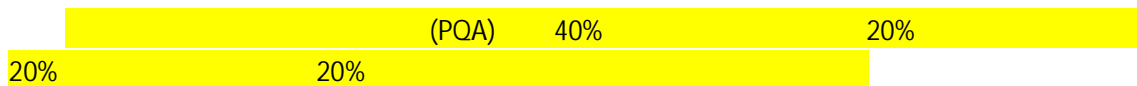
4 PSNR

$(\log_{10} rate_{Qp}, PSNR_{Qp})$

PSNR BD-rate 4 PSNR BD-rate/BD-PSNR

Excel

2



/

1.

PESQ PEAQ

2.

1.

60

2.

60

30

1.

WAV

2.

AVS

WAV

1.

A_i

C_i

D_i

i

$$R_i = A_i / C_i \tag{6.1}$$

R

$$R = \sum_{i=0}^{N-1} R_i \tag{6.2}$$

N

PEAQ/PESQ

1%

1%

PEAQ/PESQ

2

1.

2.

1

1.

JPEG

y

x

coord.jpg

1

coord.jpg

2.

.zip

30MB

1

1

Wavefront OBJ

1

2

Wavefront OBJ

mtl

2

3

PNG

1.

1

RMS

```

vn -0.8930 -0.2563 -0.3699
vn -0.8934 0.2560 -0.3691
vn -0.9668 0.2554 0.0000
# 4 vertex normals
vt 2.0000 2.0000 0.0000
vt 1.7500 2.0000 0.0000
vt 1.7500 1.9750 0.0000
vt 2.0000 1.9750 0.0000
# 4 texture cords
usemtl 01___Default
f 1/1/1 2/2/2 3/3/3
f 3/3/3 4/4/4 1/1/1
# 2 faces

```

2

mtl

```

newmtl
Ns
Ni
d Tr Alpha
Tf r g b
illum
Ka Kd Ks Ke r g b
map_Ka map_Kd

newmtl 01___Default
Ns 10.0000
Ni 1.5000
d 1.0000
Tr 0.0000
Tf 1.0000 1.0000 1.0000
illum 2
Ka 0.5882 0.5882 0.5882
Kd 0.5882 0.5882 0.5882
Ks 0.0000 0.0000 0.0000
Ke 0.0000 0.0000 0.0000
map_Ka maps\1.jpg
map_Kd maps\1.jpg

```



http://page.renren.com/gra_smartcity