# Camera Take Detection

## 1. Problem Description

**A** camera take is a series of consecutive frames taken by a camera. It can be cut into a sequence of segments and interleaved with other camera takes to form a scene which completes an event or a story in a video program. This is a common process in film editing. Figure 6 shows an example of camera take editing results from the movie "Finding Mr. Right". Each picture in the figure is a key frame selected from a shot (as illustrated in Figure 1 (a)), thus frames (a) to (h) represent consecutive shots, composing a scene. To take a closer look at the key frames, it is obvious that frames (a), (c) and (f) are from the same camera take, so are frames (b) and (e), as well as (d) and (g). Apparently, the shots from the same camera take could be grouped together and represented by one or more frames. It will highly reduce the throughput for further processing.



| | | | |
|---|---|---|---|
| (a) | (b) | (c) | (d) |
| (e) | (f) | (g) | (h) |

Figure 1. Examples of camera takes

Specifically, it takes the following four steps for camera take detection:

- **Frame difference calculation**: based on the assumption that two consecutive frames in a video shot should have high similarity in terms of visual content, the frame difference is calculated using color histogram (or raw pixel values for saving computational cost) as a measurement of similarity between two frames.

- **Shot detection**: if the frame difference is above some preset threshold, then a new shot is claimed. The selection of threshold is critical since it may cause over segmentation or down segmentation depending on the types of video programs (action, drama, etc.). To determine a proper threshold and further refine the detection results, certain constraints may apply, such as shot duration.

- **Key frame selection**: a key frame should properly represent the visual content of a shot. Without loss of generality, the last frame of a shot is selected as the key frame for later processing. **It is worth mentioning that more advanced techniques may be utilized to select (or generate) the most representative key frame(s).**
- **Camera take detection**: each detected shot (represented by a key frame) will be matched with the last shot in each detected camera take. If certain matching criterion is satisfied, then the current shot will be added to the end of the matched camera take. It is based on the assumption that a shot is most related to the one with closest temporal relationship. Initially, within a certain time period, we may assume the first shot as a camera take. The matching strategies vary from sift point matching to frame difference matching depending on various performance requirement.

## 2. Preliminary Experimental Results

To validate the effectiveness of the proposed camera take detection method, we carefully select three types of movie/TV series with different motion intensities for evaluation. The first type is "romantic", characterized by slow motion, close-up shots, and frequent camera takes interleaving. A 5-min video clip from the TV series "Missing You" is extracted for experiment. We first perform shot boundary detection, which is the foundation of camera take detection. The experimental results are shown in Table 1. As can be seen from the results, all of the shot boundaries are successfully detected. Although there are three over segmented shots, we are more interested in retrieving all the true shot boundaries.

Table 1. Shot boundary detection results.
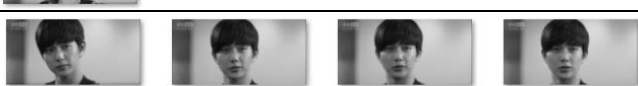
| Movie/TV Series Type | Total # Frames | Total # Shots | Detected # Shots | Precision | Recall |
|---|---|---|---|---|---|
| Romantic | 7500 | 48 | 50 | 96% | 100% |

Based on the detected shot boundaries, the camera take detection is carried out using the approach described above. The experimental results are illustrated in Table 2, where the first column represents camera take ID and the second column shows the shot array for each camera take. To be specific, each sub-picture in the shot array represents the key frame for each shot. There are totally 13 detected camera takes, which covers almost all of the detected shots. Based on the experimental results, there are several observations and conclusions: First, some camera takes are over segmented. For example, camera takes 2, 5, 6, 7, and 8 should belong to the same camera take. They are separated mostly due to the global change because of close-up effect. The

other reason would be the relatively simple background, which results in few interest points. It is also noticed that even two shots are from the same camera take, the two video sequences may not be consecutive. Second, the proposed camera take detection method is effective for romantic movies with slow-motion and little film editing (i.e., most of the shot boundaries are cut change). Third, there are some constraints for the proposed method. For example, there are about five adjusting parameters, which are empirical values and have to be tested thoroughly to adapt to different types of movies. Finally, it is worth mentioning that the total processing time is 42sec (32-bit Windows XP, 4G RAM, 2.5G Hz), which meets real-time requirement.

More experiment would be conducted for the other two types of movies (as shown in Table 3) and solid conclusions should be made accordingly.

Table 2. Camera take detection results for movie type I.

| Camera Take # | Shot Array |
|---|---|
| 1 |  |
| 2 |  |
| 3 |  |
| 4 |  |
| 5 |  |
| 6 |  |
| 7 |  |
| 8 |  |
| 9 |  |
| 10 |  |
| 11 |  |
| 12 |  |
| 13 |  |
| 14 |  |

| 15 |  | | | | |
| 16 |  | | | | |
| 17 |  | | | | |

Table 3. Summary of camera take detection results.

| Movie/TV Series Type | Total # Frames (5 min) | Detected # Shots | Detected # Camera Takes | % Stable Camera Takes | Processing Time |
|---|---|---|---|---|---|
| I. Romantic | 7500 | 50 | 13 | 100% | 38s |
| II. Low Motion (a) | 7500 | 46 | 25 | - | 46s |
| II. Low Motion (b) | 9000 | 143 | 78 | - | 142s |
| III. High Motion | 7200 | 280 | 264 | - | 143s |

## 3. Appendix

### 3.1 Camera take detection results for movie type I ("Missing You")

```
=== Camera Take 1 (13 shots)
[1, 58]
[310, 475]
[726, 855]
[1018, 1137]
[1214, 1361]
[1429, 1541]
[1639, 1673]
[1845, 1944]
[2076, 2495]
[2665, 3328]
[3459, 3514]
[3701, 3850]
[3911, 4020]

=== Camera Take 2 (8 shots)
[59, 309]
[856, 1017]
[1138, 1213]
[1362, 1428]
[1945, 2075]
[3329, 3458]
[3515, 3700]
[3851, 3910]

=== Camera Take 3 (3 shots)
[476, 519]
[520, 530]
[531, 725]
```

```
=== Camera Take 4 (1 shots)
[1542, 1638]

=== Camera Take 5 (2 shots)
[1674, 1844]
[2496, 2664]

=== Camera Take 6 (2 shots)
[4021, 4123]
[6377, 6401]

=== Camera Take 7 (7 shots)
[4124, 4203]
[4513, 4569]
[4924, 5036]
[5236, 5400]
[5707, 5777]
[6029, 6106]
[6257, 6376]

=== Camera Take 8 (2 shots)
[4204, 4259]
[4318, 4512]

=== Camera Take 9 (1 shots)
[4260, 4317]

=== Camera Take 10 (5 shots)
[4570, 4923]
[5037, 5159]
[5401, 5706]
[5778, 6028]
[6107, 6256]

=== Camera Take 11 (2 shots)
[5160, 5235]
[6473, 6530]

=== Camera Take 12 (3 shots)
[6402, 6472]
[6531, 6687]
[6823, 7327]

=== Camera Take 13 (1 shots)
[6688, 6822]
```

### 3.2 Camera take detection results for movie type II .a ("Wo Hu Cang Long")

```
=== Camera Take 1 (1 shots)
[1, 1397]

=== Camera Take 2 (1 shots)
[1398, 1499]
```

```
=== Camera Take 3 (1 shots)
[1500, 1737]

=== Camera Take 4 (1 shots)
[1738, 1883]

=== Camera Take 5 (5 shots)
[1884, 1949]
[1950, 1951]
[1952, 1953]
[1954, 1955]
[1956, 1957]

=== Camera Take 6 (1 shots)
[1958, 2152]

=== Camera Take 7 (1 shots)
[2153, 2346]

=== Camera Take 8 (1 shots)
[2347, 2635]

=== Camera Take 9 (1 shots)
[2636, 2709]

=== Camera Take 10 (2 shots)
[2710, 2810]
[2881, 3150]

=== Camera Take 11 (1 shots)
[2811, 2880]

=== Camera Take 12 (2 shots)
[3151, 3228]
[3635, 3785]

=== Camera Take 13 (1 shots)
[3229, 3363]

=== Camera Take 14 (1 shots)
[3364, 3634]

=== Camera Take 15 (3 shots)
[3786, 3919]
[4006, 4231]
[4375, 4681]

=== Camera Take 16 (1 shots)
[3920, 4005]

=== Camera Take 17 (3 shots)
[4232, 4374]
[4682, 4798]
[5112, 5191]

=== Camera Take 18 (1 shots)
[4799, 4905]
```

```
=== Camera Take 19 (2 shots)
[4906, 4968]
[5244, 5354]

=== Camera Take 20 (2 shots)
[4969, 5111]
[5192, 5243]

=== Camera Take 21 (1 shots)
[5355, 5432]

=== Camera Take 22 (1 shots)
[5433, 6107]

=== Camera Take 23 (2 shots)
[6108, 6241]
[6364, 6496]

=== Camera Take 24 (4 shots)
[6242, 6363]
[6497, 6848]
[7047, 7164]
[7260, 7476]

=== Camera Take 25 (2 shots)
[6849, 7046]
[7165, 7259]
```

### 3.3 Camera take detection results for movie type II.b ("Finding Mr. Write")

```
=== Camera Take 1 (1 shots)
[1, 73]

=== Camera Take 2 (1 shots)
[74, 166]

=== Camera Take 3 (1 shots)
[167, 331]

=== Camera Take 4 (2 shots)
[332, 494]
[721, 732]

=== Camera Take 5 (1 shots)
[495, 498]

=== Camera Take 6 (3 shots)
[499, 507]
[508, 519]
[610, 637]

=== Camera Take 7 (3 shots)
[520, 570]
```

```
[638, 673]
[674, 700]

=== Camera Take 8 (1 shots)
[571, 574]

=== Camera Take 9 (1 shots)
[575, 609]

=== Camera Take 10 (1 shots)
[701, 720]

=== Camera Take 11 (2 shots)
[733, 736]
[869, 872]

=== Camera Take 12 (1 shots)
[737, 740]

=== Camera Take 13 (1 shots)
[741, 764]

=== Camera Take 14 (1 shots)
[765, 783]

=== Camera Take 15 (1 shots)
[784, 811]

=== Camera Take 16 (1 shots)
[812, 848]

=== Camera Take 17 (1 shots)
[849, 868]

=== Camera Take 18 (1 shots)
[873, 876]

=== Camera Take 19 (1 shots)
[877, 880]

=== Camera Take 20 (1 shots)
[881, 901]

=== Camera Take 21 (1 shots)
[902, 932]

=== Camera Take 22 (1 shots)
[933, 985]

=== Camera Take 23 (1 shots)
[986, 999]

=== Camera Take 24 (1 shots)
[1000, 1012]

=== Camera Take 25 (1 shots)
[1013, 1033]
```

```
=== Camera Take 26 (1 shots)
[1034, 1052]

=== Camera Take 27 (1 shots)
[1053, 1299]

=== Camera Take 28 (1 shots)
[1300, 1335]

=== Camera Take 29 (1 shots)
[1336, 1434]

=== Camera Take 30 (1 shots)
[1435, 1520]

=== Camera Take 31 (1 shots)
[1521, 1712]

=== Camera Take 32 (1 shots)
[1713, 1808]

=== Camera Take 33 (1 shots)
[1809, 1845]

=== Camera Take 34 (14 shots)
[1846, 1851]
[1852, 1855]
[1856, 1860]
[1861, 1864]
[1865, 1868]
[1869, 1872]
[1873, 1876]
[1877, 1880]
[1881, 1884]
[1885, 1888]
[1889, 1892]
[1893, 1896]
[1897, 1900]
[1901, 1904]

=== Camera Take 35 (1 shots)
[1905, 1984]

=== Camera Take 36 (1 shots)
[1985, 2094]

=== Camera Take 37 (4 shots)
[2095, 2116]
[2117, 2122]
[2123, 2126]
[2158, 2196]

=== Camera Take 38 (3 shots)
[2127, 2157]
[2197, 2214]
[2265, 2310]
```

```
=== Camera Take 39 (1 shots)
[2215, 2264]

=== Camera Take 40 (2 shots)
[2311, 2326]
[2327, 2334]

=== Camera Take 41 (2 shots)
[2335, 2421]
[2422, 2452]

=== Camera Take 42 (1 shots)
[2453, 2468]

=== Camera Take 43 (3 shots)
[2469, 2494]
[2495, 2498]
[2499, 2534]

=== Camera Take 44 (1 shots)
[2535, 2619]

=== Camera Take 45 (2 shots)
[2620, 2671]
[2735, 2790]

=== Camera Take 46 (1 shots)
[2672, 2734]

=== Camera Take 47 (13 shots)
[2791, 2829]
[2855, 2950]
[3008, 3064]
[3121, 3229]
[3280, 3334]
[3370, 3510]
[3568, 3638]
[3717, 3854]
[3961, 4144]
[4183, 4222]
[4268, 4426]
[4483, 4514]
[4553, 4804]

=== Camera Take 48 (7 shots)
[2830, 2854]
[3065, 3120]
[3335, 3369]
[3855, 3960]
[4145, 4182]
[4427, 4482]
[4515, 4552]

=== Camera Take 49 (5 shots)
[2951, 3007]
[3230, 3279]
```

```
[3511, 3567]
[3639, 3716]
[4223, 4267]

=== Camera Take 50 (1 shots)
[4805, 4834]

=== Camera Take 51 (1 shots)
[4835, 4911]

=== Camera Take 52 (5 shots)
[4912, 5067]
[5248, 5329]
[7462, 7495]
[7582, 7756]
[8078, 8139]

=== Camera Take 53 (1 shots)
[5068, 5145]

=== Camera Take 54 (1 shots)
[5146, 5188]

=== Camera Take 55 (1 shots)
[5189, 5247]

=== Camera Take 56 (1 shots)
[5330, 5766]

=== Camera Take 57 (1 shots)
[5767, 5809]

=== Camera Take 58 (1 shots)
[5810, 6258]

=== Camera Take 59 (1 shots)
[6259, 6340]

=== Camera Take 60 (1 shots)
[6341, 6614]

=== Camera Take 61 (1 shots)
[6615, 6799]

=== Camera Take 62 (2 shots)
[6800, 6877]
[6953, 7250]

=== Camera Take 63 (1 shots)
[6878, 6952]

=== Camera Take 64 (1 shots)
[7251, 7314]

=== Camera Take 65 (1 shots)
[7315, 7377]
```

```
=== Camera Take 66 (1 shots)
[7378, 7416]

=== Camera Take 67 (1 shots)
[7417, 7461]

=== Camera Take 68 (2 shots)
[7496, 7581]
[7803, 7876]

=== Camera Take 69 (2 shots)
[7757, 7802]
[7877, 7921]

=== Camera Take 70 (2 shots)
[7922, 7940]
[7941, 7952]

=== Camera Take 71 (1 shots)
[7953, 8077]

=== Camera Take 72 (3 shots)
[8140, 8247]
[8355, 8396]
[8503, 8546]

=== Camera Take 73 (1 shots)
[8248, 8354]

=== Camera Take 74 (1 shots)
[8397, 8502]

=== Camera Take 75 (4 shots)
[8547, 8581]
[8641, 8666]
[8727, 8832]
[8905, 8979]

=== Camera Take 76 (1 shots)
[8582, 8600]

=== Camera Take 77 (1 shots)
[8601, 8640]

=== Camera Take 78 (2 shots)
[8667, 8726]
[8833, 8904]
```