Free-viewpoint Video Generation for Sports Scene

KDDI R&D Laboratories, Inc.

Hiroshi SANKOH

July 8, 2014

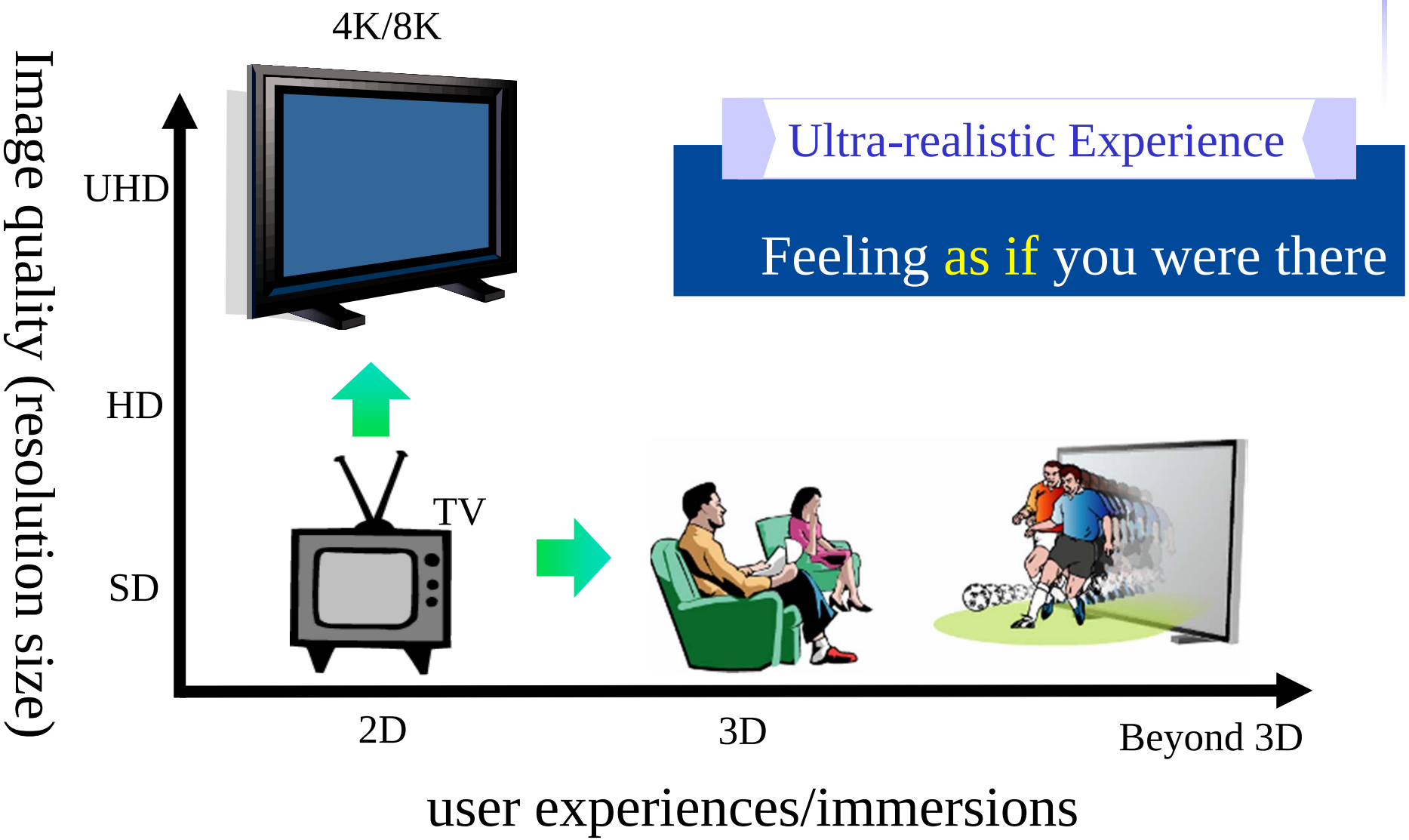
Outlines

1. Introduction
   - Overview of the existing practical applications and research works

2. Details of generation methods
   - Technical issues for sports scene
   - Our proposed algorithm

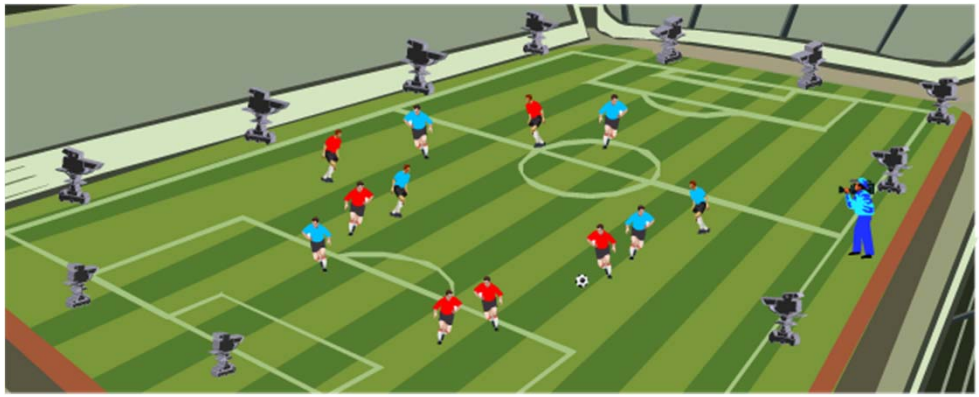
3. Conclusion and Future works
   - Toward 2020 Tokyo Olympics/Paralympics

Ultra-realistic Image media



Toward Beyond 3D

- Free-viewpoint 3D video can realize:
  - your own favorite angles for real event videos. (Basic function in CG or video game)
  - virtual viewpoint where cameras cannot be set. (ex. Viewpoint of a goalkeeper)
  - stereoscopic viewing based on both eye's disparity



Shooting of multi-view videos

Virtual viewpoint

Practical Applications

KDDI R&D LABS

- Bullet-time
  - Shots with 120 calibrated cameras
  - The Matrix (1999)

- Eyevision
  - Controls pan and tilt parameters of 32 cameras
  - Super Bowl XXXV (2001)

Based on Prof. Kanade's ``virtualized reality" activity

Practical Applications

KDDI R&D LABS

- LiberoVision (commercial product)
  - Interpolating 2 camera images for outdoor sports

Many of the existing practical applications take not multiple frames but only still frame as input

Related Research Works

KDDI R&D LABS

■Representative synthesis methods

- Model-based approach (3D model reconstruction)
- Image-based approach (Interpolation, Ray space method)

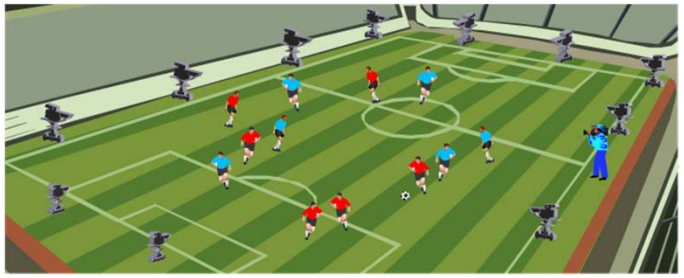
-> selected by the requirements for virtual viewpoint

■Requirements for virtual viewpoint

- Dynamic view-change: Walk-through and Fly-through:

  —passing by multiple objects and

  -> Model-based method is appropriate



Viewpoint Selection

Two types of Modeling methods

KDDI R&D LABS

■Accurate model (Polygon mesh)

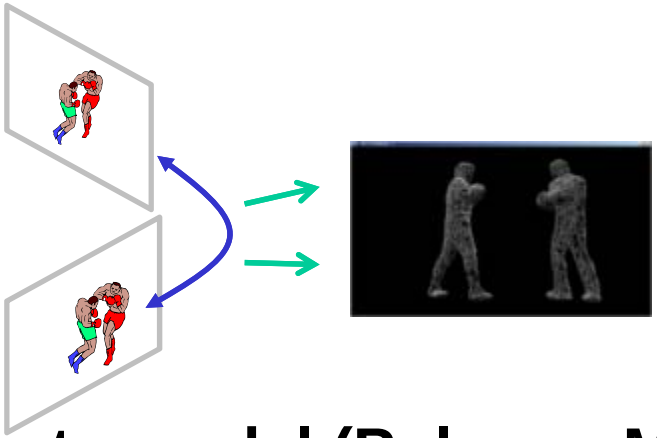
- Merit: naturally reproduces view-changing from arbitrary viewpoints
- Drawback: A number of calibrated cameras are necessary

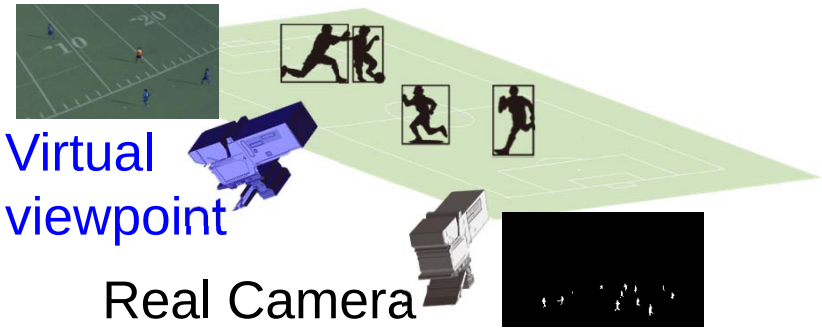
-> appropriate for indoor space such as 5 meters by 5 meters

■Simple model (Billboard)

- Merit: applicable for at least a single non-calibrated camera
- Drawback: a minute view-changing cannot be realized

-> suitable for wide outdoor space like soccer stadium





Accurate model (Polygon Mesh) **Simple model (Billboard)**

Demo (1/2)

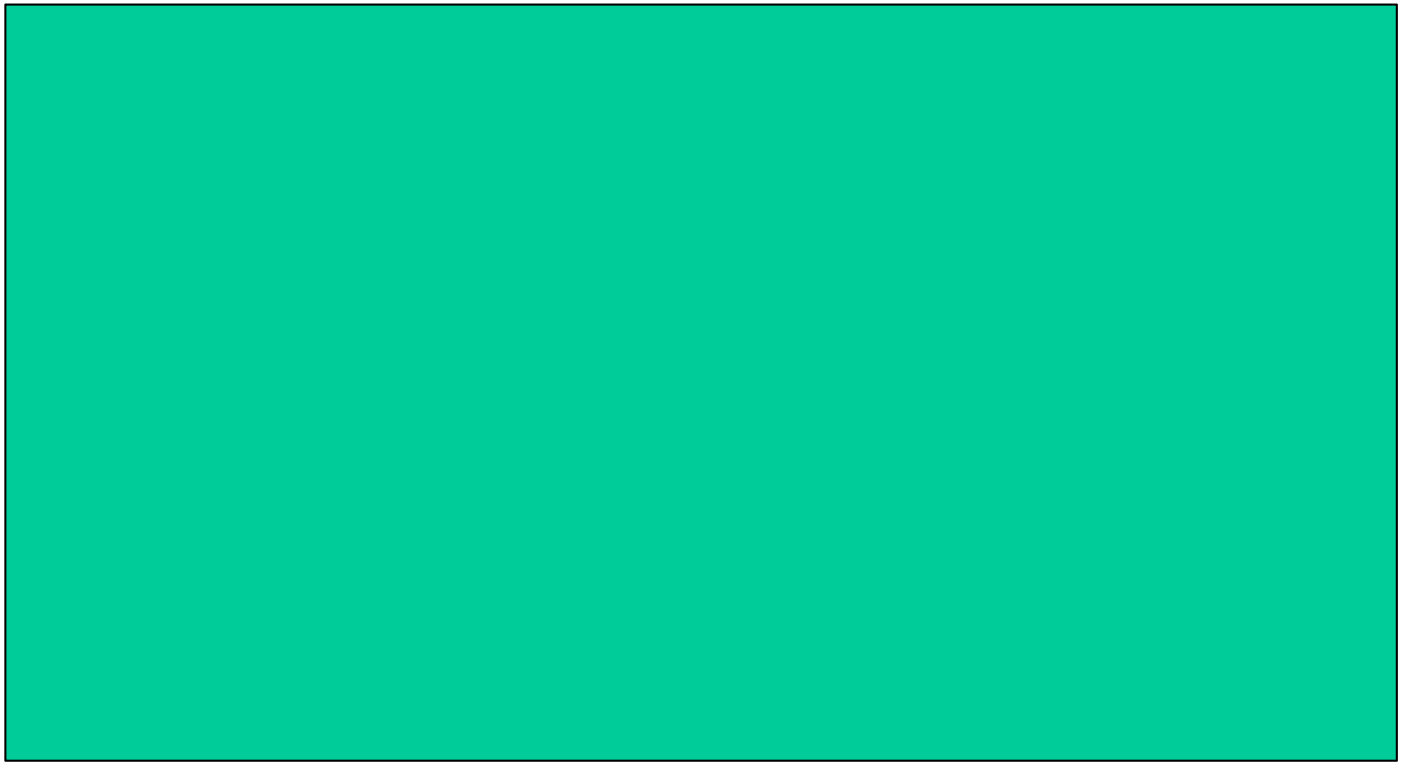
- Polygon Mesh reconstructed from 30 HD cameras
  - Walk-through and Fly-through experiences

Used in Final presentation for Bid of FIFA World-cup 2022

Demo (2/2)

KDDI R&D LABS

- Billboard represented by a single 4K camera



Exhibited at CEATEC JAPAN 2010, CES 2011

Outlines

1. Introduction
   - Overview of the existing practical applications and research works

2. Details of generation methods
   - Technical issues for sports scene
   - Our proposed algorithm

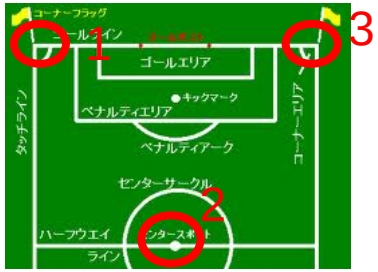
3. Conclusion and Future works
   - Toward 2020 Tokyo Olympics/Paralympics

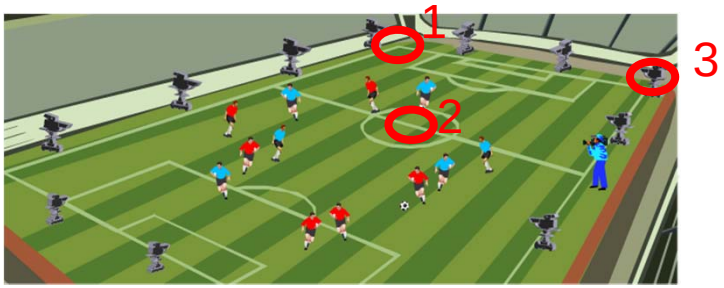
Billboard based free-viewpoint video

KDDI R&D LABS

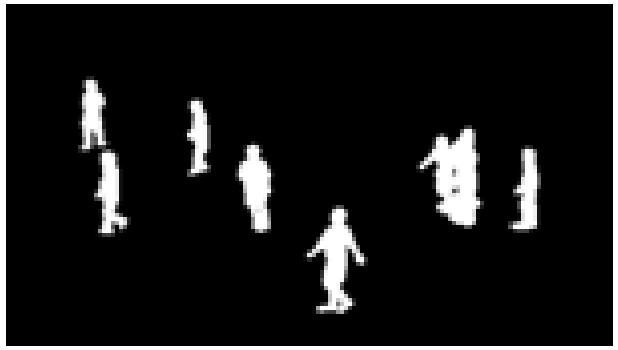
■Procedures of the synthesis

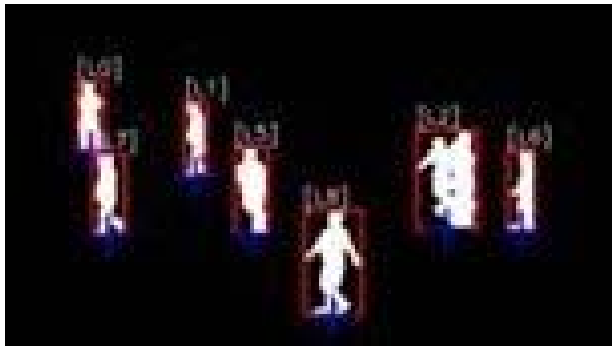
1. Estimation of Homograph matrix
   — World coordinate of the field plane is standardized





2. Extraction of silhouette region for each object





Billboard based free-viewpoint video

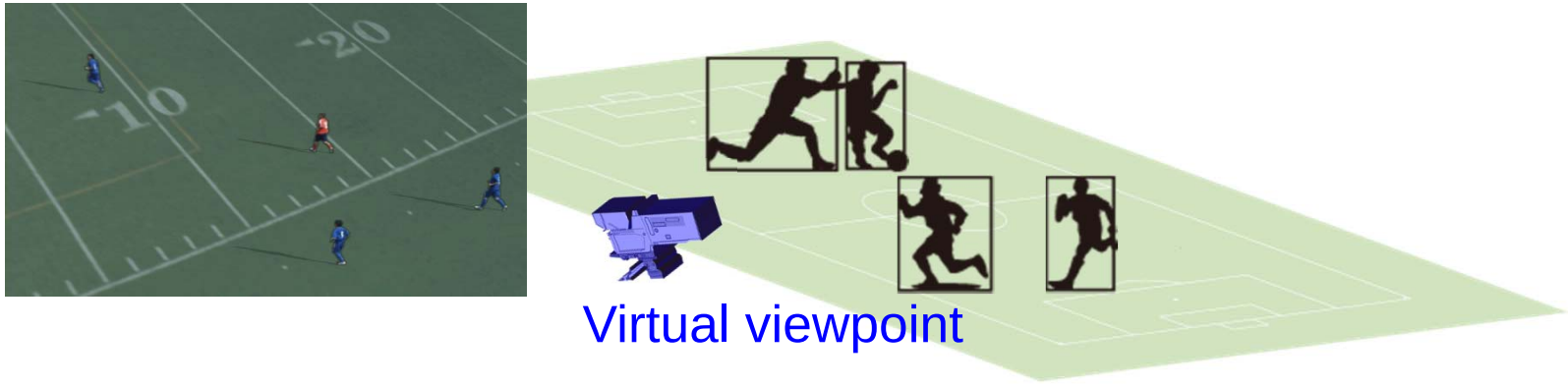
KDDI R&D LABS

■Procedures of the synthesis

3. Estimating the world coordinate of each object's position

—Calculated with homograph matrix and silhouette regions

4. Synthesizes virtual viewpoint videos



Real camera

Technical issues for billboard method

KDDI R&D LABS

■Camera movements and Occlusions

1. Moving cameras (only fixed cameras are treated in conventional)

—Camera parameters (Homograph matrix) changes at every frame

—Extraction of each object is difficult (Background subtraction)

-> Cannnot be applied to sports live broadcasting

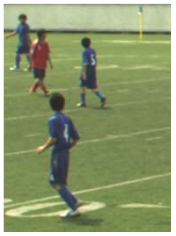
2. <u>Occlusion handling for multiple objects</u>

—Lack of textures for occluded objects

—Difficulty for estimation for object's positions

Frame 44

Frame 56



Frame 64

Frame 76

Artifacts caused by occlusion problems

KDDI R&D LABS

■Original captured video sequence

Artifacts caused by occlusion problems

KDDI R&D LABS

■Synthesized Video (Bird-view from the sky)

- Some objects' Positions frequently change

Overview of the proposed algorithm

KDDI R&D LABS

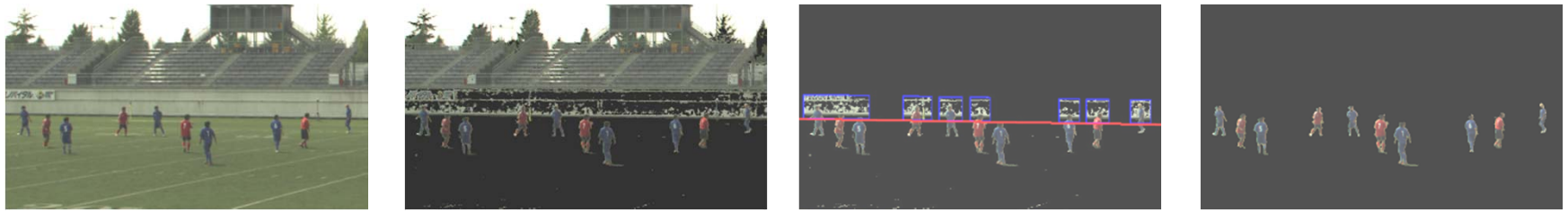
■Dynamic Camera Calibration

- Correspond points between frame t and t-1
  - —Surf features robust for image rotation and zoom-in and out
  - -> suitable for pan-tilt-zoom camera used in broadcasting

Frame t-1 Frame t

Frame t

- Object extraction based on estimated homograph matrix
  - —Border line detection between the field plane and spectator areas



Overview of the proposed algorithm

KDDI R&D LABS

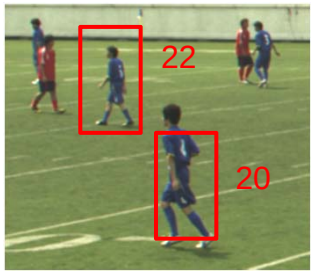
■Object Tracking

- Modify Textures and world coordinate positions for occlusion regions

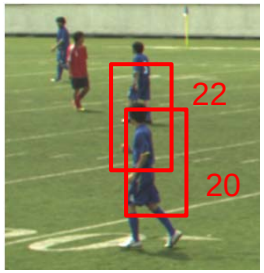
  —Tracks each object with the consistent IDs for consecutive frames among multiple cameras

  -> Detects occlusion regions for every frame

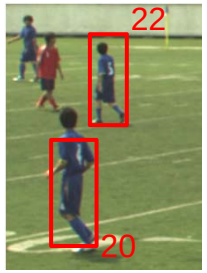
  -> interpolate the world coordinate and texture between frames



Frame 44



Frame 56



Frame 64

Effectiveness of the proposed method

KDDI R&D LABS

■Synthesized video based on the proposed method

- Flicker movements of some players are interpolated

Synthesized free-viewpoint video (1/2)

KDDI R&D LABS

- Captured with a single 4K moving camera
  - All Japan University Soccer Championship 2013 (Final Match)

Synthesized free-viewpoint video (2/2)

KDDI R&D LABS

■Captured with a single 4K moving camera

- All Japan University Soccer Championship 2013 (Final Match)

Outlines

1. Introduction
   - Overview of the existing practical applications and research works

2. Details of generation methods
   - Technical issues for sports scene
   - Our proposed algorithm

3. Conclusion and Future works
   - Toward 2020 Tokyo Olympics/Paralympics

Conclusions

KDDI R&D LABS

- Free-viewpoint video generation for sports scene based on billboard model was summarized.
  - Details of generation method
    - —Dynamic camera calibration
    - —Occlusion interpolation based on object tracking
    - -> Additional improvements are necessary to realize perfectly automatic processing
  - Recent technical trials
    - —All Japan University Soccer Championship 2013
    - —The championship series in NPB 2013

Toward 2020 Tokyo Olympics/Paralympics

Real time generation of

Free-viewpoint video

Player centric viewpoint

Live broadcasting of real-time event for general users

For general users' sports scene such as golf and dance and so on, we would like to provide this technology with users

Requirements for FTV standard (Draft)

KDDI R&D LABS

- Data format of our billboard based method
  - Identified number for each object consistent among multiple frames
  - Silhouette (including alpha) and Texture of each object
  - Billboard model (World coordinate and direction) on the ground plane
  - Space (Background) CG Model
  - Camera parameter for CG Space