

만화책 형식 동작 데이터 시각화

이강훈

광운대학교 컴퓨터소프트웨어학과

kang@kw.ac.kr

Visualizing Motion Data as Sequential Images in Comic Book Layout

Kang Hoon Lee

Dept. of Computer Science, Kwangwoon University

요약

오늘날 실제 사람의 움직임을 촬영하여 얻은 동작 데이터는 영화 제작, 게임 개발, 재활 의학, 스포츠 분석 등의 다양한 분야에서 활용되고 있다. 시간에 따라 변화하는 일련의 자세로 이루어진 동작 데이터는 그 내용을 신속하고 정확하게 파악하는 것이 쉽지 않다. 이는 특히 많은 양의 동작 데이터를 관리하고 원하는 동작을 찾고자 할 때 장애가 되곤 한다. 널리 사용되는 한 가지 방법은 동영상을 재생하듯이 관찰 시점을 진전시키는 동시에 3차원 공간 상의 시야를 조절하는 것이다. 하지만 이 방식은 시점과 시야의 두 요소를 함께 고려하면서 데이터의 길이에 비례하는 시간 동안 지속적인 상호작용을 수행하여야 하는 번거로움을 야기한다. 본 논문은 동작 데이터를 분석하여 일련의 효과적인 관찰 시점 및 시야를 자동으로 찾아내고, 이로부터 생성된 이미지 시퀀스를 만화책 레이아웃 형식으로 배치하는 새로운 시각화 방법을 제안한다. 사용자는 만화책을 읽는 것처럼 빠르게 전체 흐름을 파악할 수 있고, 원하는 동작이 등장하는 부분을 쉽게 찾아낼 수 있다. 이동, 권투, 환경과의 상호작용 등 여러 종류의 동작 데이터를 만화책 형식으로 시각화한 결과를 보임으로써 제안된 방법의 유용성을 확인하였다.

Abstract

Captured motion data is widely used today in a variety of areas including film production, game development, sports analysis, and medical rehabilitation. The ability of analyzing and processing motion data has increased rapidly for last decades. However, it is still difficult for users to quickly understand the contents of motion data consisting of a series of time-varying poses. One typical approach is to visualize consecutive poses in sequence while adjusting three-dimensional view, which is often time-consuming and laborious especially when users need to repeatedly control time and view in order to search for desired motions. We present a method of visualizing motion data as a sequence of images in comic book layout so that users can rapidly understand the overall flows of motion data, and easily identify their desired motions. The usefulness of our approach is demonstrated by visualizing various kinds of motion data including locomotion, boxing, and interaction with environments.

키워드: 동작 데이터 요약, 만화책 형식 레이아웃, 동작 분할, 시야 선택, 비사실적 렌더링

Keywords: motion data summary, comic book layout, motion segmentation, view selection, non-photorealistic rendering

1. 서론

모션 캡처 시스템의 성능이 향상되고 비용이 감소하면서 영화 제작, 게임 개발, 재활 의학, 스포츠 분석 등의 다양한 목적을 위한 인간 동작 데이터의 수요가 증가하는 추세이다. 많은

수의 동작 클립을 사용하는 환경에서 각 클립에 포함된 동작의 내용과 종류를 신속하게 파악하는 것은 매우 중요하다. 동작 클립은 시간에 따라 변화하는 일련의 자세 집합으로 구성된다. 이를 시각화하는 가장 기본적인 방법은 동영상을 재생

하듯이 순차적으로 자세를 렌더링하는 것이다. 사용자가 원하는 대로 구도를 조절할 수 있고, 필요할 경우 구간을 이동할 수도 있다는 점에서 이 방법은 높은 융통성을 제공한다. 하지만 동작 클립의 길이가 길 수록 전체 내용을 확인하는데 오랜 시간이 걸린다는 한계가 있다.

전체 동작의 내용을 신속하게 전달하는 방법으로서 모든 자세를 동시에 하나의 이미지에 렌더링하는 방법을 생각할 수 있다. 이 방법은 매우 높은 압축률을 제공하지만, 실제로 활용하기 어려운 정도의 시각적 정보 손실을 야기한다. 이미지 공간에서 많은 수의 자세가 중첩되어 각각의 자세를 식별하거나 자세 간의 선후 관계를 파악하기가 어렵기 때문이다. 모든 자세를 렌더링하는 대신 일부 자세만을 선별하여 표현하면 자세 식별 및 동작 흐름의 파악을 보다 용이하게 할 수 있다. 특히 상대적으로 중요한 자세들을 우선적으로 선별하면, 단순히 규칙적인 간격으로 추출하는 것에 비하여 효과적인 요약이 가능하다 [1]. 하지만 제자리에서 다양한 물건을 취하거나 긴 시간 동안 제한된 영역 내에서 이동하는 등의 동작이 주어지면 내용을 상당 부분 생략하지 않는 한 자세 간 중첩을 피하는 것이 매우 어려워진다.

요약 이미지의 표현 복잡도를 줄이는 한 가지 방법은 두 개 이상의 순차적인 이미지로 분할하여 병치하는 것이다. 각 이미지는 전체 동작 데이터의 일부 구간만을 요약하기 때문에 공간적 중첩을 줄일 수 있다. 이와 같이 다수의 이미지를 병렬적으로 배치하는 형식은 만화 장르의 주요 특성 중 하나로서, 시간 순서에 따른 이야기나 정보를 간결하게 축약하도록 도와주는 중요한 도구이다. 본 논문은 동작 데이터를 다수의 구간으로 분할하여 각 구간마다 별도의 이미지로 표현한 후 이들 이미지를 만화책 레이아웃으로 배치하는 새로운 시각화 방식을 제안한다 (그림 1).

사용자가 칸의 개수, 행의 개수, 쪽의 너비와 높이 정보를 제공하면 입력 동작 데이터는 자동적으로 네 단계의 과정을 거쳐 만화책 레이아웃의 이미지로 변환된다. 먼저 동작의 변화가 감지되는 일련의 주요 시점을 선별하여 시간적으로 축약된 동작 데이터를 구성한다 (3장). 축약된 동작 데이터를 여러 칸에 걸쳐 표현하기 위하여 사용자가 지정한 개수 만큼의 구간으로 분할하고 (4장), 각 구간에 포함된 자세들을 가장 효과적으로 관찰할 수 있는 시점 및 시야를 결정한다 (5장). 끝으로, 각 구간마다 결정된 시점과 시야에 따라 렌더링을 수행하여 칸에 해당하는 이미지를 생성하고, 이들 칸을 사용자가 지정한 행의 개수에 맞추어 배치하여 쪽 전체의 이미지를 합성한다 (6장).

본 논문이 제안하는 시각화 방법은 각 구간을 명료하게 표현하는 동시에 전체 흐름을 신속하게 전달하는 장점을 제공한다. 이를 실험적으로 확인하기 위하여 Carnegie Mellon 대학 라이브러리에서 발췌한 일련의 이동 동작 [2], 제자리에서 취하는 다양한 물체 동작, 복잡한 환경과 상호작용하는 동작 등 다양한 종류의 데이터를 시각화하였다 (7장).

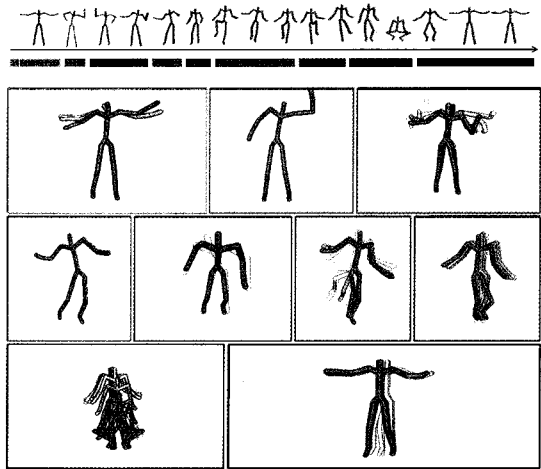


그림 1: 만화책 형식의 시각화.

2. 관련 연구

모션 캡처 기술의 발전과 수요 증가에 따라 이를 통해 확보한 인간 동작 데이터의 내용을 효율적으로 탐색하고 기술하기 위한 여러 가지 분석 방법이 개발되었다. 자세 간 유사도 분석과 클러스터링 [3, 4], 차수 감소 기법을 통한 2차원 공간 상의 자세 배열 [5], 행위 변화에 따른 동작 구간 분할 [6], 매개화 가능한 유사 동작 탐색 [7, 8], 동작 패턴의 반복성과 연결성 분석 [9] 등 동작 데이터에 내재되어 있는 구조를 파악하고 시각화하는 다양한 접근방식은 동작 생성과 편집 기술의 중요한 밑바탕이 되고 있다.

일반 사용자의 입장에서 동작 데이터의 내용을 쉽게 파악할 수 있도록 하는 효과적인 시각화 방법에 관한 연구는 비교적 최근에 이르러 활발히 이루어지고 있다. 동작 데이터의 시각화에는 크게 관찰 시점과 구도의 두 가지 요소가 개입된다. 비디오 클립을 재생하듯이 처음부터 마지막까지 모든 순간을 차례대로 관찰할 경우에는, 주요 움직임이 최대한 정확하게 포착할 수 있는 최적의 카메라 궤적을 찾아내는 문제로 귀결된다 [10]. 반면, 동일한 구도에 여러 시점의 자세를 동시에 표현함으로써 하나의 이미지로 신속하게 전체 내용을 전달하는 것도 가능하다. Assa 등은 동작 곡선의 극점에 위치한 일련의 주요 자세를 식별한 후, 이들 자세의 공간적 분포가 최대화되는 구도를 선택하여 렌더링함으로써 비교적 짧은 시간 동안 이루어지는 행위를 간결하게 묘사할 수 있음을 보였다 [1]. 추가적으로 화살표, 물결 모양, 고속촬영 효과 등의 동세 표현기법을 이용하여 정지 자세에 동적인 느낌을 부여할 수도 있다 [11].

동작 데이터가 시간에 따른 정지 자세의 집합이라면, 비디오 데이터는 시간에 따른 정지 영상의 집합으로서 한 번에 전체 내용을 파악하기 어렵다는 문제점을 공유한다. 한편, 동작

데이터에 비하여 상대적으로 높은 접근성 덕분에 비디오 데이터의 신속한 열람을 위한 표현 방법의 연구는 ‘비디오 요약’이라는 주제로 HCI, 멀티미디어, 컴퓨터 비전, 컴퓨터 그래픽스 등의 분야에서 오랜 기간 많은 주목을 받아왔다 [12]. 비디오 요약 연구의 접근방법은 그 결과로 산출되는 미디어의 종류에 따라 분류되곤 한다. 일부 연구 결과는 입력 비디오의 연속된 프레임 구간을 제거하거나 서로 결합함으로써 보다 짧게 축약된 비디오를 출력으로 생성한다 [13, 14]. 반면에 입력 비디오로부터 일련의 키프레임을 선택한 후 해당 이미지들을 모자이크 형식으로 잘라 붙이거나 병치함으로써 하나의 개요 이미지를 합성하는 방식도 활발히 연구되고 있다 [15, 16]. 특히, 본 논문의 접근방식과 유사하게 만화책 레이아웃의 형태로 키프레임 이미지를 배치하는 방식도 알려져 있다 [17, 18].

3. 동작 축약

동작 데이터는 시간에 따라 변화하는 일련의 자세 집합 $\{x_t | 1 \leq t \leq F\}$ 로 구성되고, 각 자세는 공간 상의 위치 \mathbf{p}_t 와 신체 골격을 구성하는 관절들의 회전 방향 $\{\mathbf{q}_{t,j} | 1 \leq j \leq J\}$ 에 의하여 표현된다. 특정 자세를 이미지 형태로 시각화하기 위해서는 먼저 관절의 회전 정보로부터 전체 골격의 형태를 계산하고, 이에 부합하는 다각형 모델을 구성하여 해당 위치에 배치한 후 3차원 렌더링을 수행한다.

단일 자세의 시각화는 짧은 노출 시간을 설정한 상태에서 사진을 촬영하는 것과 유사하다. 빠르게 달리는 자동차를 촬영한 경우에도 마치 정지 상태에서 촬영한 것처럼 보이듯이, 하나의 자세만을 렌더링하였을 때 움직임에 대한 정보는 상실된다. 사진가들은 이와 같은 제약을 극복하고자 연속촬영 기법을 사용하곤 한다. 짧은 시간 간격으로 포착된 여러 순간의 이미지를 하나의 필름 안에 합성하여 움직임을 표현하는 방식이다. 동작 데이터에 속한 일련의 자세를 추출하여 하나의 이미지에 동시에 렌더링하면 이와 유사한 효과를 얻을 수 있다. 이때 단순히 일정 시간 간격으로 자세를 추출하는 대신, 동작의 변화를 포착하는 주요 자세만을 추출할 경우 보다 적은 개수의 자세를 이용하여 효과적으로 동작을 표현할 수 있다.

손, 발 등 골격구조 말단부의 가속도 크기가 극대, 극소에 이르는 시점은 신체 동작의 변화를 가리키는 주요 단서를 제공한다. 걸거나 뛸 때 발이 땅에 닿는 시점, 점프할 때 최고점에 도달하거나 착지하는 시점, 주먹을 뺏거나 발차기를 할 때 가속이 이루어지는 시점 등이 여기에 속한다. 하지만 이들 시점은 지나치게 미세한 변화들을 포착하기 때문에 본 연구에서는 그 중 자세 변화가 일정 범위 이상으로 커지는 시점만을 선별한다. 서로 다른 두 자세 \mathbf{x}_t 와 $\mathbf{x}_{t'}$ 간의 차이는 다음과 같이 Lee 등 [4]의 방법을 사용하여 계산한다.

$$D(\mathbf{x}_t, \mathbf{x}_{t'}) = \|\mathbf{p}_t - \mathbf{p}_{t'}\|^2 + \sum_{j=1}^J \omega_j \|\log(\mathbf{q}_{t,j}^{-1} \mathbf{q}_{t',j})\|^2 \quad (1)$$

여기에서 ω_j 는 해당 관절의 회전이 골격 형태에 미치는 영향을 고려하여 실험적으로 결정하는 가중치이다. 가속도 크기가 극점에 도달하는 모든 시점을 나열한 후, 순차적으로 각 시점을 탐색하면서 가장 최근 시점과의 자세 차이가 임계값 이하이면 제거하고, 임계값 이상일 경우에만 선택하는 과정을 반복한다. 이 과정을 통하여 선택된 자세들의 집합은 일종의 ‘축약된’ 동작 데이터로 볼 수 있고, 본 논문에서는 원래 데이터와의 구분을 위하여 $\{y_t | 1 \leq t \leq f\}$ 로 표기한다. 자세 차이의 임계값을 d 라고 하면, 이 동작 데이터는 항상 $D(y_t, y_{t+1}) > d$ 를 만족한다. 이후의 장에서는 $\{y_t\}$ 에 포함된 각각의 자세 및 변화의 흐름을 명료하게 전달하기 위한 시각화 과정을 기술한다.

4. 구간 분할

축약된 동작 데이터의 전체 구간 $[1, f]$ 가 주어졌을 때, 이를 n 개의 세부 구간 $t_1 = [s_1, e_1], t_2 = [s_2, e_2], \dots, t_n = [s_n, e_n]$ ($s_1 = 1, e_n = f, s_i \leq e_i = s_{i+1} - 1$)으로 나누어 각 구간에 포함된 자세를 별도의 칸에 시각화하고자 한다. 구간의 개수 n 을 줄일수록 전체 동작이 적은 개수의 칸에 압축되어 빠른 시간 안에 전달하는 것이 가능하지만, 자세가 복잡하게 얽혀 명확한 전달이 어려워질 수 있다. 반면 n 을 늘릴수록 각 칸에서 표현하는 자세의 개수가 줄어들기 때문에 이미지의 복잡도는 감소하지만 전체 흐름을 파악하는데 걸리는 시간은 늘어날 것이다 (그림 2). 사용자는 n 을 조절하여 원하는 수준만큼 간략하게 혹은 상세하게 묘사된 결과를 얻을 수 있다.

동일한 개수로 분할하더라도 어느 시점에서 분할하느냐에 따라서 이미지의 복잡도가 달라질 수 있다. 예를 들어 한 방향으로 걷다가 어느 순간 반대 방향으로 되돌아 오는 동작을 측면에서 관찰한다고 하자. 방향을 전환하는 시점에서 분할하면 각각의 이미지를 자세 중첩 없이 명료하게 표현할 수 있겠지만, 그 밖의 시점에서 분할하면 전진과 후진 동작의 일부 영역이 중첩되어 이미지가 복잡하게 표현될 것이다.

최적의 분할 시점을 결정하기 위해서는 임의의 구간 $t = [s, e]$ 에 대한 이미지 복잡도를 측정할 수 있어야 한다. 하지만 아직 관찰 시야가 결정되지 않은 상태이므로, 현 단계에서는 시야와 상관 없이 주어진 구간의 위치 및 자세 정보만을 이용하여 다음과 같이 근사적으로 추정하는 방식을 사용한다.

- **이동 밀집도:** 동일한 개수의 자세를 표현할 경우, 제자리에서 몸짓을 취하는 동작은 이동하는 동작보다 복잡할 가능성이 크다. 또한 동일한 속도로 이동할 경우, 여러 방향으로 오가는 동작이 일정한 방향으로 이동하는 동작보다

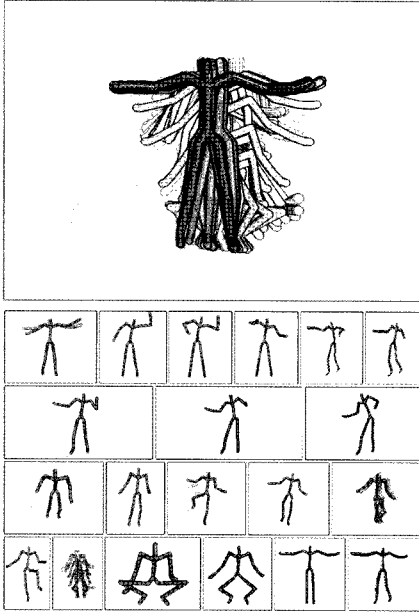


그림 2: (상) 1개 구간으로 압축. (하) 20개 구간으로 압축.

복잡할 가능성이 크다. 이는 주된 이동 방향을 기준으로 측정된 이동 범위가 좁을 수록 이미지 상에서 복잡하게 표현될 것임을 의미한다. 주된 이동 방향을 구하기 위하여 주어진 구간의 모든 위치 좌표 $\{p_s, \dots, p_e\}$ 에 대하여 주성분 분석(principal component analysis)을 수행한다. 가장 분산이 큰 방향의 축으로 위치 좌표를 사영하여 대응되는 1차원 좌표 집합 $\{x_s, \dots, x_e\}$ 를 얻으면, 이로부터 주된 이동 방향에서의 범위 $[x_{min}, x_{max}]$ 를 찾을 수 있다. 구간 t 에 포함된 자세의 개수를 $\Delta t (= e - s + 1)$ 로 표현하면 단위 구간에 대한 이동 범위는 $\frac{x_{max} - x_{min}}{\Delta t}$ 로 나타낼 수 있고, 이동 밀집도는 그 역수로 측정한다.

- **자세 다양성:** 동일 구간에 서로 유사한 자세가 많을 수록 보다 적은 개수의 자세로 요약하여 이미지 복잡도를 줄일 수 있는 여지가 커진다. 반면, 자세의 다양성이 커질 수록 하나의 장면 안에 명료하게 표현하기 어려워질 것이다. 자세의 다양성을 측정하기 위하여 먼저 k -평균 클러스터링 방법을 이용하여 전체 구간의 자세 집합 $\{y_t\}$ 를 수식 (1)의 거리 함수에 따라 일정 개수의 클러스터로 분류한다. 일단 분류가 이루어지면, 임의의 구간에 대한 자세 다양성은 단위 구간에 등장하는 클러스터의 개수로 측정할 수 있다. 즉 t 번째 자세가 속한 클러스터의 인덱스를 I_t 라고 하면, 구간 t 에 대한 자세 다양성은 $\frac{|I_s, \dots, I_e|}{\Delta t}$ 로 측정한다.

이미지 복잡도는 이동 밀집도와 자세 다양성을 동시에 고려하기 위하여 두 값을 곱한 결과로 추정한다. 이는 주어진 구간에 포함된 자세의 개수 Δt 를 상쇄시켜 다음의 식을 유도한다.

$$C(t) = \frac{|I_s, \dots, I_e|}{x_{max} - x_{min}} \quad (2)$$

최적의 분할 (t_1^*, \dots, t_n^*) 은 가능한 모든 분할 중에서 다음과 같이 전체 이미지 복잡도의 총 합계를 최소화하는 경우로 정의할 수 있다:

$$(t_1^*, \dots, t_n^*) = \underset{(t_1, \dots, t_n) \in T_n}{\operatorname{arg\,min}} \sum_{i=1}^n C(t_i) \quad (3)$$

여기에서 T_n 은 주어진 구간을 n 개로 나누는 가능한 모든 분할의 집합을 나타낸다. 동작 데이터의 길이 f 와 분할 개수 n 이 증가할 수록, 경우의 수 $|T_n|$ 이 급격하게 증가하기 때문에 실용적인 시간 범위 내에서 전역 최적해를 찾는 것은 매우 어렵다. 대신 본 논문은 반복적 구간 병합을 통하여 분할 시점을 찾는 알고리즘을 제안한다. 기본 아이디어는 전체 구간을 충분히 많은 개수의 세부 구간으로 잘게 나눈 후, n 개의 구간이 남을 때까지 복잡도 증가를 최소화하는 인접 구간을 찾아 반복적으로 병합하는 것이다. 특정 순간 m 개의 구간이 남아있다고 가정하면, 다음과 같이 병합 후의 복잡도와 병합 전의 평균 복잡도 간의 차를 최소화하는 인접 구간 t_i 와 t_{i+1} 을 찾은 후 하나의 새로운 구간 $t'_i = [s_i, e_{i+1}]$ 으로 병합한다.

$$i = \underset{j \in \{1, 2, \dots, m-1\}}{\operatorname{arg\,min}} \left(C(t_j) - \frac{C(t_j) + C(t_{j+1})}{2} \right) \quad (4)$$

5. 시점과 시야 결정

시간적으로 분할된 각 구간의 동작은 운동 방향, 이동 범위, 시간 간격 등이 서로 다르기 때문에 구간 별로 해당 동작을 효과적으로 표현할 수 있는 시점과 시야를 결정할 필요가 있다. 잘못된 시야를 선택하면 구간에 포함된 일부 혹은 전체 자세가 배제될 수도 있고, 자세 간의 중첩이 심하여 각각의 자세나 동작의 흐름을 파악하기 어려울 수도 있다. 이 장에서는 자세 간의 중첩을 최소화하는 시야를 자동으로 찾는 방법을 제시한다. 또한 여러 구간에 걸쳐 진행되는 동작의 흐름을 쉽게 파악하도록 하기 위하여 인접한 구간 간의 시야 차이를 최소화하고자 한다.

컴퓨터 그래픽스에서 시야는 일반적으로 관찰 위치 e , 방향 $(\vec{u}, \vec{v}, \vec{n})$, 범위 $(\theta_u, \theta_v, z_n, z_f)$ 에 의하여 결정된다 (그림 3). 원근 시야의 경우, 3차원 공간에 배치된 관찰 대상은 위치 e 를 중심으로 이로부터 z_n 만큼 떨어진 위치에 놓인 가상의 이미

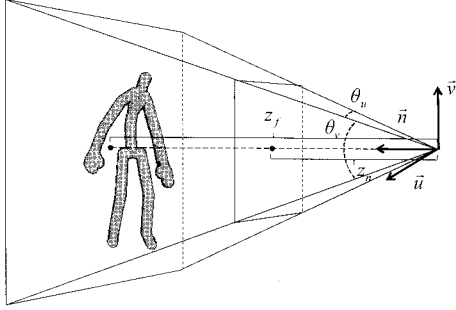


그림 3: 시야를 결정하는 일련의 변수들.

지 평면에 투영되어 2차원 이미지를 형성한다. 이미지 평면은 \vec{n} 에 수직하고, \vec{u} 와 \vec{v} 에 의하여 그 좌표계가 정의된다. 이미지 평면의 가로 방향 \vec{u} 에 해당되는 각도 범위 θ_u , 세로 방향 \vec{v} 에 해당되는 각도 범위 θ_v , 그리고 관찰 위치로부터 이미지 평면에 수직인 방향 \vec{n} 의 거리 범위 $[z_n, z_f]$ 는 시야의 3차원 영역을 규정하며, 이 영역을 벗어난 관찰 대상은 투영 과정에서 제외된다. 이들 변수가 모두 독립적이라고 가정하면 높은 차원으로 인하여 최적의 시야를 찾기 위한 탐색 공간의 크기가 매우 커진다. 하지만 그 중 많은 부분은 대상을 전혀 관찰하지 않거나, 지나치게 멀리서 혹은 가까이에서 관찰하는 등 실제적으로 거의 의미가 없는 시야에 해당된다.

본 연구에서는 시야 변수에 대하여 몇 가지 제약조건을 부과함으로써 대상을 적절하게 포착하는 시야로 탐색 공간을 한정한다. 이미지 평면의 수평 방향 \vec{u} 는 항상 전역 좌표계의 x - z 평면, 즉 수평면과 평행을 유지하도록 한다. 만화에서는 연출 기법의 하나로 기울어진 시야를 사용하기도 하지만, 동작 데이터의 시각화에 있어서는 거의 필요성이 없다. 평행 조건을 만족시키기 위하여 관찰 방향을 전역 좌표계에 대한 x -축 중심 회전 각도 ϕ 와 y -축 중심 회전 각도 θ 로 다음과 같이 매개화한다.

$$\begin{aligned} \vec{u} &= (\cos \theta, 0, -\sin \theta) \\ \vec{v} &= (-\sin \phi \sin \theta, \cos \phi, -\sin \phi \cos \theta) \\ \vec{n} &= (-\cos \phi \sin \theta, -\sin \phi, -\cos \phi \cos \theta) \end{aligned} \quad (5)$$

시야의 세로 방향 범위 θ_v 는 고정하고 가로 방향 범위 θ_u 만 변화할 수 있도록 한다. 이는 최종 레이아웃을 구성할 때 칸의 높이는 고정하고 너비만 변화할 수 있도록 하는 것과 대응된다 (6장). 시야의 중횡비와 칸의 중횡비를 최대한 동일하게 유지함으로써 투영으로 인한 왜곡 정도의 변화를 최소화할 수 있다. 실험에서 θ_v 는 $\pi/4$ 라디안으로 고정하였다.

시선이 관찰 대상의 중심을 바라보도록 하고, 관찰 대상이 시야 영역을 거의 가득 차지하도록 제약하여 나머지 모든 변

수를 유도할 수 있다. 효율적인 계산을 위하여 먼저 관찰 대상을 시야 방향 (\vec{u} , \vec{v} , \vec{n})에 따라 정렬된 최소 경계 상자로 근사한다. 주어진 구간에 포함된 모든 자세를 감싸는 최소 경계 상자의 위치를 \mathbf{c} , 각 방향 크기를 (s_u, s_v, s_n)라고 하면, 대상의 중심을 바라보는 관찰 위치는 다음과 같이 표현할 수 있다.

$$\mathbf{e} = \mathbf{c} - d \cdot \vec{n} \quad (6)$$

대상의 중심으로부터 관찰 위치까지의 거리 d 는 경계 상자가 시야의 세로 방향 범위를 벗어나지 않는 한도에서 최소한의 값으로 결정한다.

$$d = \tan\left(\frac{\pi}{2} - \frac{\theta_v}{2}\right) \cdot \frac{s_v}{2} + \frac{s_n}{2} \quad (7)$$

일단 관찰 위치가 결정되면, 시야의 시선 방향 범위 $[z_n, z_f]$ 는 각각 경계 상자를 포함하는 최대, 최소의 값으로서 $[d - s_n/2, d + s_n/2]$ 로 결정한다. 또한 시야의 가로 방향 범위 θ_u 는 역시 경계 상자를 포함하는 최소값으로서 다음과 같이 계산한다.

$$\theta_u = 2 \cdot \tan^{-1}\left(\frac{s_u}{2d - s_n}\right) \quad (8)$$

상수로 정의된 θ_v 를 제외한 모든 시야 변수를 두 개의 변수 ϕ 와 θ 로부터 유도하기 때문에 본래의 차원보다 현격히 낮은 2차원 공간에서 효과적으로 최적의 시야를 탐색할 수 있다. 본 연구에서는 일정 범위 안에서 균일하게 추출한 (ϕ, θ) 쌍의 가능한 모든 조합에 대하여 시야의 품질을 평가하고, 그 중 가장 높은 품질을 제공하는 시야를 선택한다. θ 는 이전 구간에서 선택된 θ_{prev} 로부터 크게 벗어나지 않도록 하기 위하여 $\theta_{prev} - \frac{\pi}{2} \leq \theta \leq \theta_{prev} + \frac{\pi}{2}$ 로 범위를 제한하고, 첫번째 구간에 한하여 전체 범위인 $-\pi \leq \theta \leq \pi$ 를 허용한다. 또한 아래에서 위로 쳐다보는 시야, 혹은 지나치게 아래로 내려다보는 시야를 배제하기 위하여 ϕ 의 범위는 $-\frac{\pi}{4} \leq \phi \leq 0$ 으로 한정한다. 본 연구의 실험에서는 ϕ 와 θ 모두 이와 같은 범위 안에서 10도 간격으로 일정하게 추출하였다.

임의의 시야 (ϕ, θ) 에 대한 품질은 이전 구간에서 선택된 시야와의 차이, 이미지 평면 상에서 자세의 분산도와 가시도, 그리고 깊이 순서의 정렬도에 의하여 측정한다.

$$Q = \omega_d Q_d + \omega_s Q_s + \omega_v Q_v + \omega_a Q_a \quad (9)$$

$\omega_d, \omega_s, \omega_v, \omega_a$ 는 모두 0 이상의 실수이고 그 합이 1인 가중치이다. 서로 인접한 구간에서 시선 방향이 급격히 변화하면 동작의 연결 관계를 이해하기 어려워진다. Q_d 는 이 같은 상황을

방지하기 위하여 이전 구간에서 선택된 시선 방향 \vec{n}_p 와 현재 시선 방향 \vec{n} 이 서로 일치하는 정도를 측정한다.

$$Q_d = \vec{n} \cdot \vec{n}_p \quad (10)$$

자세들이 이미지 평면 상에서 넓게 분포될 수록 각 자세의 식별 및 전체 흐름의 파악이 용이해진다. 각각의 자세 y_i 를 골격 주요 부위에 해당하는 점들의 집합 $\{\mathbf{p}_{t,i}\}_{i=1}^m$ 으로 단순화하고 전체 자세에 대하여 주성분 분석을 수행하였을 때, 가장 큰 고유치를 갖는 두 개의 축 \vec{e}_x, \vec{e}_y 는 투영된 점들의 분포를 최대화하는 평면을 형성한다. 이는 이들 두 축에 수직인 세 번째 축 \vec{e}_z 방향으로 관찰할 때 이미지 평면 상에서 자세들이 가장 넓게 분포될 것임을 의미한다. Q_s 는 시선 방향이 이 축과 얼마나 가까운지 계산하여 자세 분포도를 추정한다.

$$Q_s = \vec{n} \cdot \vec{e}_z \quad (11)$$

자세 가시도 Q_v 는 이미지 평면 상에서 각 자세가 다른 자세와의 중첩 없이 얼마나 잘 표현되는지 측정한다. 점 집합 $\{\mathbf{p}_{t,i}\}_{i=1}^m$ 으로부터 골격 구조에 따라서 인접한 두 점을 연결한 선분들의 집합을 $\{l_{t,j}\}_{j=1}^n$ 이라고 하면, 임의의 두 자세 간 중첩도는 이미지 평면에 투영된 모든 선분 쌍에 대한 교차 검사를 통하여 계산할 수 있다.

$$O(t, t') = \frac{1}{n^2} \sum_{j=1}^n \sum_{k=1}^n \text{INT}(T(l_{t,j}), T(l_{t',k})) \quad (12)$$

여기에서 T 는 시야 (ϕ, θ) 에 의하여 3차원 공간 상에 배치된 임의의 선분을 이미지 평면 상에 투영하는 변환을 나타내고, INT는 두 선분이 교차할 경우 1, 교차하지 않을 경우 0을 반환하는 함수이다. 결과적으로 $O(t, t')$ 은 $[0, 1]$ 의 범위를 갖게 되고, 두 자세가 이미지 평면 상에서 전혀 교차하지 않으면 0, 모든 선분 쌍이 교차하면 1을 반환한다. 주어진 구간의 모든 자세 쌍에 대하여 이와 같은 교차 검사를 수행하면 계산 복잡도가 지나치게 커진다. 대신 Q_v 는 인접 구간과의 동작 연결을 전달하는데 중요한 역할을 수행하는 처음 자세와 마지막 자세에 대하여서만 다른 모든 자세와의 교차 검사를 수행하여 가시도를 측정한다.

$$Q_v = 2(\Delta t - 1) - \left(\sum_{t=s+1}^e O(s, t) + \sum_{t=s}^{e-1} O(t, e) \right) \quad (13)$$

깊이 순서의 정렬도 Q_a 는 자세의 시간 순서와 시야 좌표계 상에서의 깊이 방향 순서가 어느 정도 일치하는지 측정한다. 모든 인접한 두 자세 쌍에 대하여 나중 자세가 이전 자세보다

시점에 가까울 수록 정렬이 잘 되어 있다고 판단한다.

$$Q_a = \frac{1}{\Delta t - 1} \sum_{t=s}^{e-1} D(t, t+1) \quad (14)$$

여기에서 $D(i, j)$ 는 시야 (ϕ, θ) 좌표계 상에서 i 번째 자세의 최소 경계 상자 깊이 범위 $[d_i^{min}, d_i^{max}]$ 와 j 번째 자세의 최소 경계 상자 깊이 범위 $[d_j^{min}, d_j^{max}]$ 를 비교하여, $d_i^{min} < d_j^{min}$ 과 $d_i^{max} < d_j^{max}$ 를 모두 만족할 때 1, 그렇지 않을 때 0을 반환하는 함수이다. 6장에서 설명하듯이, 각 구간의 자세를 렌더링 할 때 항상 나중 자세는 이전 자세를 덮어 쓰도록 하기 때문에 Q_a 가 클 수록 깊이에 대한 왜곡이 적어지고 자연스러운 결과를 보여주게 된다.

6. 시각화

본 논문에서는 전체 동작 데이터를 만화책의 한 쪽에 해당하는 하나의 큰 사각형 영역 안에 시각화한다. 시간적 분할을 전혀 하지 않았을 경우, 모든 자세를 꼭 전체 영역에 동시에 표현한다. 두 개 이상의 구간으로 분할하였다면, 전체 영역을 구간 개수만큼의 작은 칸들로 나누는 다음 각 칸마다 대응되는 구간의 자세들을 표현한다. 일련의 구간은 일반적으로 만화책을 읽는 순서인 왼쪽에서 오른쪽, 위에서 아래 방향의 흐름에 따라 일련의 칸에 순차적으로 대응된다.

만화책에서 하나의 쪽을 여러 개의 칸으로 나누는 방식, 즉 레이아웃 형식은 규칙적인 격자 형태부터 서로 다른 크기와 모양의 칸들로 구성된 불규칙적인 형태에 이르기까지 매우 다양하다. 본 연구에서는 모든 칸의 높이를 고정시키고 너비 변화만을 허용하는 절충 형태의 레이아웃을 사용한다. 이와 같은 레이아웃은 쉽게 구현이 가능하고 계산 효율성이 높은 동시에, 완전히 규칙적인 격자 형태에 비해서는 장면의 특성에 좀 더 유연하게 대응할 수 있다는 장점을 제공한다. 하지만 하나의 칸이 여러 행을 차지하거나 직각이 아닌 형태를 갖도록 허용하는 등 보다 폭넓은 범위에서 효과적인 레이아웃을 모색하는 것은 향후 흥미로운 연구 과제로 남아있다.

사용자가 쪽의 너비 w 와 높이 h , 그리고 행의 개수 r 를 지정하면 자동으로 전체 레이아웃이 구성된다. 너비 변화만을 허용하는 제약에 따라 모든 칸의 높이 h_i 는 동일하게 h/r 이 된다. 행의 개수 r 는 미리 분할된 구간의 개수, 즉 표현하여야 할 칸의 개수 n 보다 작거나 같다고 가정한다. 이상적으로 각 칸의 너비 w_i 는 시야 선택시 계산된 중형비를 유지하기 위하여 $\hat{w}_i = h_i \cdot \tan \theta_u$ 와 같아야 하지만, 하나 이상의 칸이 고정된 너비의 행 안에 배치되어야 하므로 이를 보장할 수 없다. 대신 각 칸의 실제 너비는 시야 중형비에 비례하여 행의 너비 w 를 분배하도록 한다. 즉, 임의의 행 \mathbf{r} 에 $\{t_p, t_{p+1}, \dots, t_q\}$ 의 구간이 포함될 경우 각 칸의 너비 $w_i (p \leq i \leq q)$ 는 다음과 같이 결정한다

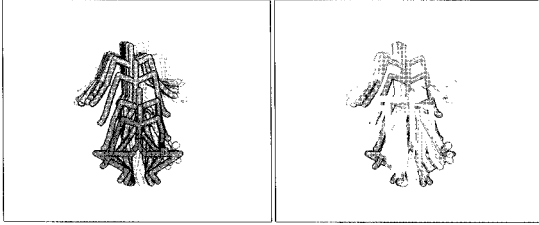


그림 4: (좌) 순서 우선 렌더링. (우) 깊이 우선 렌더링.

다.

$$w_i = w \cdot \frac{\hat{w}_i}{\sum_{j=p}^q \hat{w}_j} \quad (15)$$

칸의 실제 너비 w_i 와 이상적인 너비 \hat{w}_i 간의 차이가 클수록 표현하고자 하는 정보를 잃거나 왜곡하는 정도가 커진다. 본 연구에서는 가능하면 이와 같은 정보 손실 및 왜곡을 최소화하는 방향으로 칸을 나누고자 한다. 이를 위하여 다음과 같이 행 복잡도를 정의한다.

$$\tilde{C}(r) = \sum_{j=p}^q (\hat{w}_j - w_j)^2 \quad (16)$$

n 개의 칸을 r 개의 행에 분배할 때는 4장에서 반복적으로 인접 구간을 병합하는 방식과 유사하게, 처음에는 n 개의 행에 하나씩 칸을 분배한 후 r 개의 행만 남을 때까지 복잡도 증가를 최소화하는 인접 행을 선택하여 반복적으로 병합한다.

레이아웃 구성이 완료되면, 모든 구간에 대하여 미리 결정된 시야에 따라 3차원 렌더링을 수행하여 대응되는 칸 영역에 결과 이미지를 배치한다. 여기에서 선택적으로 렌더링 전에 시야 중첩비와 칸 중첩비가 일치되도록 $\theta_u = \tan^{-1}(w_i/h_i \cdot \tan \theta_v)$ 로 조정할 수 있다. 이는 최종 확정된 칸 너비 w_i 가 이상적인 너비 \hat{w}_i 보다 작을 경우, 원래 관찰하고자 하는 대상의 일부가 칸 바깥으로 배제될 것임을 의미한다. 반면 이와 같이 조정하지 않을 경우 모든 대상이 칸 안에 표현되기는 하지만 이미지가 세로 방향으로 길게 늘려진 모습으로 왜곡된다.

각 구간의 자세는 시간 순서에 따라 렌더링하고, 항상 다음 자세는 이전 모든 자세를 덮어 쓰도록 하여 3차원 상의 깊이 정보보다 시간 순서에 대한 정보를 우선적으로 제공한다. 그림 4에서 확인할 수 있듯이, 깊이 정보를 우선할 경우 자세 서로 간의 중첩으로 인하여 각각의 자세뿐 아니라 동작의 흐름을 나타내는 자세의 순서까지 식별하기 어려워진다. 반면 시간 정보를 우선할 경우, 5장에서 깊이 방향 정렬이 잘 이루어진 시야를 선택하였기 때문에 대부분의 경우 동작의 흐름을 명확하게 전달할 수 있다. 또한 자세를 칠하는 색상의 밝기를

점점 어둡게 함으로써 순서에 대한 추가적인 시각적 단서를 제공한다.

7. 실험 결과

본 논문이 제안하는 시각화 방법의 유용성을 확인하기 위하여 여러 종류의 동작 데이터에 대하여 다양한 크기와 상세도의 요약본을 생성하였다. 각각의 칸을 너비 2048, 높이 2048 크기의 텍스처에 미리 렌더링한 후 너비 1240, 높이 877의 화면 프레임버퍼에 복사하여 배치하는 방식을 사용하였다. 화면에는 한 번에 전체 쪽 영역의 일부에 속하는 3개의 행만 시각화하였고, 사용자는 대화형 인터페이스를 통하여 이전 혹은 다음 영역으로 이동하며 결과를 확인할 수 있다. 이는 마치 여러 쪽으로 구성된 만화책을 읽는 것과 유사한 경험을 제공한다.

- **이동.** Carnegie Mellon 대학 동작 데이터베이스에서 발췌한 일련의 이동 동작 집합에 적용하여 보았다(그림 5(a)). 각각의 클립은 수 초의 분량으로 짧은 편이어서 주요 자세의 개수는 10~20개 정도이다. 또한 그 중 대부분의 자세가 이동 과정에 포함되어 있고, 제자리에서 취하는 몸짓은 극소수이기 때문에 3~5개의 칸 만으로도 효과적인 축약이 가능하다. 이동 양상이 변하는 시점에 맞추어 적절하게 시간 분할이 이루어졌고, 이동 방향에 대하여 수직에 가까운 시야가 자동적으로 선택됨으로써 동세를 선명하게 드러냄을 확인할 수 있다.
- **권투.** 권투 선수가 미리 주어진 시나리오 없이 가상의 상대를 가정하고 펀치와 풋웍의 다양한 조합을 훈련하는 1분 30초 분량의 동작 데이터를 가시화하였다. 실험 결과 대략 50개 정도의 구간으로 나누었을 때 효과적인 시각적 축약 결과를 얻을 수 있었다. 단순한 이동 동작 데이터와 달리 제자리에서 펀치 동작을 수행하는 구간과 풋웍을 통하여 자리를 옮기는 구간이 번갈아 나타난다. 그림 5(b)에서 볼 수 있듯이, 이동하는 구간은 비교적 긴 시간에 걸친 동작을 하나의 칸으로 압축하여 묘사하고, 제자리에서 펀치를 수행하는 구간은 짧은 시간 간격으로 분리하여 자세 변화를 중첩 없이 표현하고 있다.
- **놀이터.** 미끄럼틀과 계단으로 구성된 작은 놀이터 구조물에서 돌아다니고, 기어오르고, 미끄럼틀을 타고 내려오는 등의 동작을 반복한 10분 분량의 동작 데이터를 시각화하였다(그림 5(c)). 수평 방향의 이동뿐 아니라 높이 방향의 이동까지 함께 등장하기 때문에 깊이 방향 정렬도가 평균적으로 낮은 편이다. 결과적으로 자세의 시간적 순서는 명확히 전달하지만 상대적인 위치 관계를 왜곡하는 경우가 몇 차례 등장한다. 이와 같은 문제점을 해결할 수 있는 보다 효과적인 렌더링 방식에 대한 추가적인 연구가 요구된다.

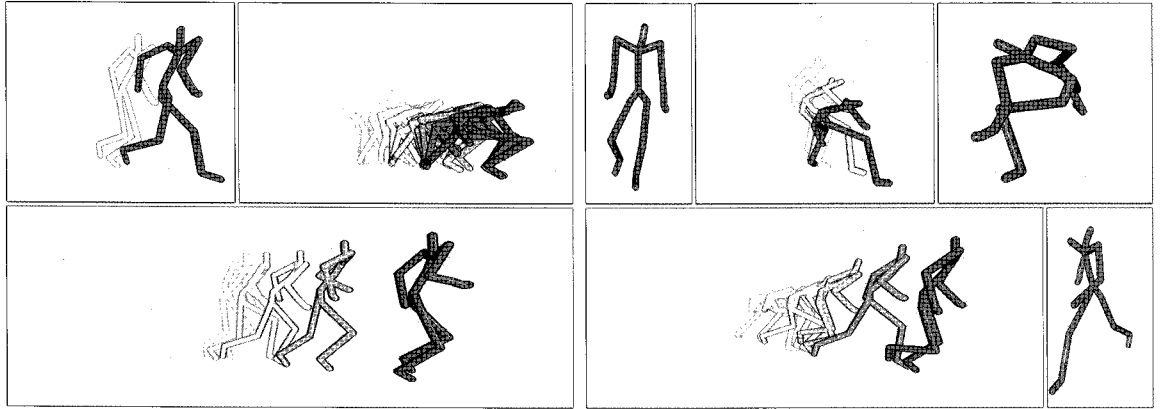
8. 결론

본 논문은 동작 데이터를 축약이 용이한 일련의 구간으로 분할하고 각 구간에 대하여 효과적인 시야를 자동으로 선택하여 렌더링한 뒤, 그 결과로 생성된 일련의 이미지를 만화책 레이아웃에 맞춰 배치하는 새로운 시각화 방법을 제안하였다. 또한 이동, 권투, 구조물 사용 등 여러 종류의 동작 데이터에 적용한 시각화 결과를 제시함으로써 본 논문이 제안하는 방법의 유효성을 보였다.

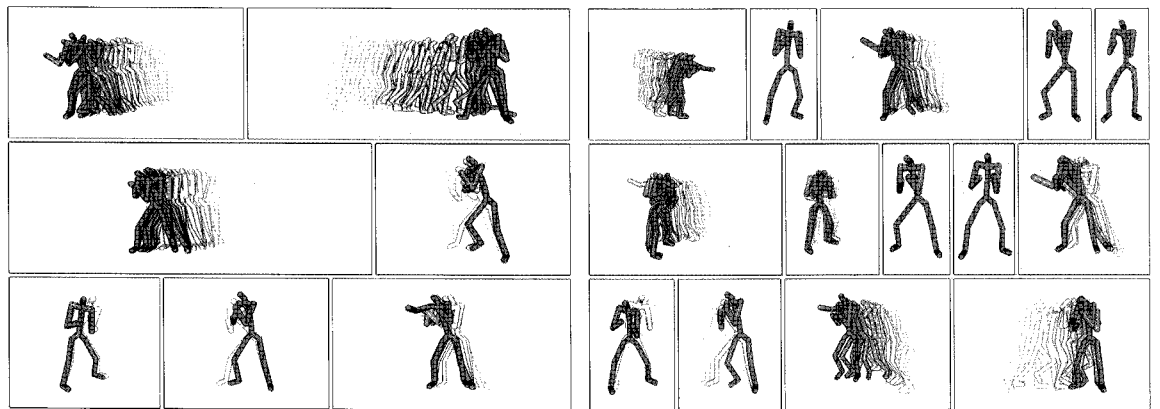
본 연구는 동작 데이터의 효과적인 사용자 인터페이스를 탐구하는 초기 단계 결과의 하나로 볼 수 있다. 향후 보다 정밀한 시각적 압축 방식에 대한 연구뿐 아니라, 이와 같은 새로운 시각화 방식에 기초한 편리한 데이터 접근과 활용 인터페이스에 대한 연구 역시 활발히 이루어질 것으로 기대한다.

참고 문헌

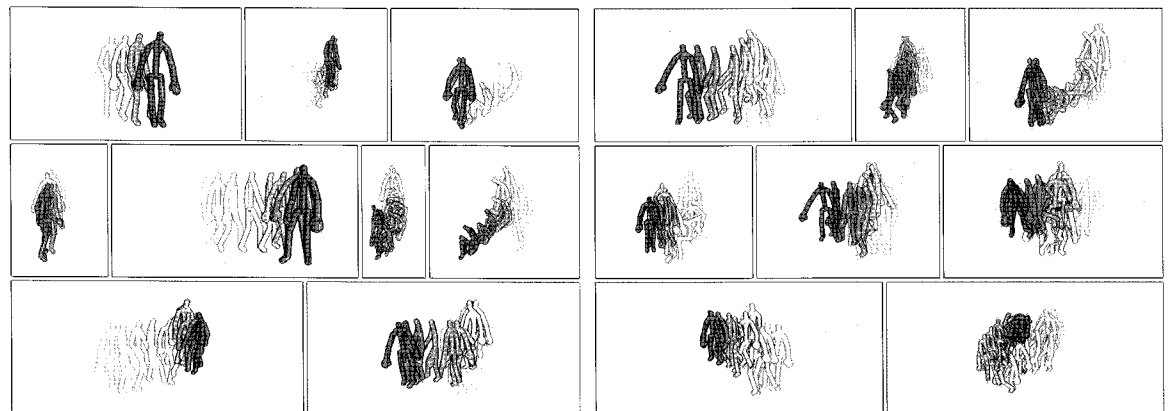
- [1] J. Assa, Y. Caspi, and D. Cohen-Or, "Action synopsis: pose selection and illustration," *ACM Transactions on Graphics (SIGGRAPH 2005)*, vol. 24, no. 3, pp. 667–676, 2005.
- [2] CMU, "Graphics laboratory motion capture database," <http://mocap.cs.cmu.edu>.
- [3] L. Kovar, M. Gleicher, and F. Pighin, "Motion graphs," *ACM Transactions on Graphics (SIGGRAPH 2002)*, vol. 21, no. 3, pp. 473–482, 2002.
- [4] J. Lee, J. Chai, P. S. A. Reitsma, J. K. Hodgins, and N. S. Pollard, "Interactive control of avatars animated with human motion data," *ACM Transactions on Graphics (SIGGRAPH 2002)*, vol. 21, no. 3, pp. 491–500, 2002.
- [5] Y. Sakamoto, S. Kuriyama, and T. Kaneko, "Motion map: image-based retrieval and segmentation of motion data," in *SCA '04: Proceedings of the 2004 ACM SIGGRAPH/Eurographics symposium on Computer animation*, 2004, pp. 259–266.
- [6] J. Barbič, A. Safonova, J.-Y. Pan, C. Faloutsos, J. K. Hodgins, and N. S. Pollard, "Segmenting motion capture data into distinct behaviors," in *GI '04: Proceedings of Graphics Interface 2004*, 2004, pp. 185–194.
- [7] L. Kovar and M. Gleicher, "Automated extraction and parameterization of motions in large data sets," *ACM Transactions on Graphics (SIGGRAPH 2004)*, vol. 23, no. 3, pp. 559–568, 2004.
- [8] H. J. Shin and H. S. Oh, "Fat graphs: constructing an interactive character with continuous controls," in *SCA '06: Proceedings of the 2006 ACM SIGGRAPH/Eurographics symposium on Computer animation*, 2006, pp. 291–298.
- [9] P. Beaudoin, M. van de Panne, P. Poulin, and S. Coros, "Motion-motif graphs," in *SCA '08: Proceedings of the 2006 ACM SIGGRAPH/Eurographics symposium on Computer animation*, 2008, pp. 117–126.
- [10] J. Assa, D. Cohen-Or, I.-C. Yeh, and T.-Y. Lee, "Motion overview of human actions," *ACM Transactions on Graphics (SIGGRAPH ASIA 2008)*, vol. 27, no. 5, pp. 1–10, 2008.
- [11] S. Bouvier-Zappa, V. Ostromoukhov, and P. Poulin, "Motion cues for illustration of skeletal motion capture data," in *NPAR '07: Proceedings of the 5th international symposium on Non-photorealistic animation and rendering*, 2007, pp. 133–140.
- [12] B. T. Truong and S. Venkatesh, "Video abstraction: A systematic review and classification," *ACM Transactions on Multimedia Computing, Communication and Applications*, vol. 3, no. 1, p. 3, 2007.
- [13] E. P. Bennett and L. McMillan, "Computational time-lapse video," *ACM Transactions on Graphics (SIGGRAPH 2007)*, vol. 26, no. 3, p. 102, 2007.
- [14] Y. Pritch, A. Rav-Acha, and S. Peleg, "Nonchronological video synopsis and indexing," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 11, pp. 1971–1984, 2008.
- [15] M. Irani and P. Anandan, "Video indexing based on mosaic representations," *Proceedings of the IEEE*, vol. 86, no. 5, pp. 905–921, 1998.
- [16] D. B. Goldman, B. Curless, D. Salesin, and S. M. Seitz, "Schematic storyboarding for video visualization and editing," *ACM Transactions on Graphics (SIGGRAPH 2006)*, vol. 25, no. 3, pp. 862–871, 2006.
- [17] J. Boreczky, A. Girgensohn, G. Golovchinsky, and S. Uchihashi, "An interactive comic book presentation for exploring video," in *CHI '00: Proceedings of the SIGCHI conference on Human factors in computing systems*, 2000, pp. 185–192.
- [18] J. Čalić, D. P. Gibson, and N. W. Campbell, "Efficient layout of comic-like video summaries," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 17, no. 7, pp. 931–936, 2007.



(a) 이동



(b) 권투



(c) 뛰어

그림 5: 다양한 동작데이터의 시각화.

〈저자 소개〉



이강훈

- 2000년 2월 서울대학교 컴퓨터공학과 (학사)
- 2002년 2월 서울대학교 컴퓨터공학부 (석사)
- 2007년 8월 서울대학교 컴퓨터공학부 (박사)
- 2007년 10월 ~ 2008년 2월 서울대학교 컴퓨터공학부 박사후연구원
- 2008년 3월 ~ 현재 광운대학교 컴퓨터소프트웨어학과 조교수
- 〈관심분야〉 컴퓨터 그래픽스, 캐릭터 애니메이션, 컴퓨터 게임, 인간-컴퓨터 상호작용 등