

Design and Implementation of a Boundary Matching System Supporting Partial Denoising for Large Image Databases

Bum-Soo Kim*, Jin-Uk Kim*

Abstract

In this paper, we design and implement a partial denoising boundary matching system using indexing techniques. Converting boundary images to time-series makes it feasible to perform a fast search using indexes even on a very large image database. Thus, using this converting method we develop a client-server system based on the previous partial denoising research in the GUI(graphical user interface) environment. The client first converts a query image given by a user to a time-series and sends denoising parameters and the tolerance with this time-series to the server. The server identifies similar images from the index by evaluating a range query, which is constructed using inputs given from the client and sends the resulting images to the client. Experimental results show that our system provides many intuitive and accurate matching results.

▶ Keyword: Boundary matching, Indexing, Partial denoising, Time-series matching, Data mining

1. Introduction

최근 데이터 크기의 기하급수적인 증가로 인해 데이터베이스에서 유용한 정보를 재빨리 찾을 수 있는 방법들에 대한 절실한 필요성이 대두되고 있다. 최근에는 대용량 데이터베이스를 대상으로 윤곽선 매칭 문제를 해결하는 연구들이 있어 왔다[1, 2, 3, 4]. 그림 1은 윤곽선 매칭에 대한 다양한 연구들의 예제이다. 그림을 보면, 그림 1(a)는 스케일-불변 윤곽선 매칭에서 윤곽선 스케일링 수행 예시를 나타내며, 그림 1(b)는 전체 노이즈 제거 윤곽선 매칭에서 전체 노이즈 제거하는 과정을 그림으로 표현하였다. 그림 1(c)는 회전-불변 윤곽선 매칭에서 윤곽선을 회전하여 질의 윤곽선과 비교하는 예시를 표현하였으며, 그림 1(d)는 본 논문에서 다루는 부분 노이즈제거 윤곽선 매칭에서 발생할 수 있는 부분 노이즈 예시들을 나타냈다. 비록, 본 논문에서 다루는 윤곽선 매칭 방법은 폐곡선의 오브젝트 이미지에 적합한 방법이지만, 그림에서와 같이 윤곽선 이미지를 시계열로 변환하면 대용량의 이미지 데이터베이스라 할지라도 수월하게 윤곽선 매칭이 가능하다. 여기에서 시계열(time-series)이란, 각 시간에 따라 측정할 실수 값의 시퀀스이

다. 윤곽선 매칭은 먼저 윤곽선 이미지를 시계열로 변환한 후, 주어진 질의 시계열과 유사한 데이터 시계열들을 판별하는 작업을 의미한다. 여기서, 윤곽선 이미지가 변환된 시계열을 윤곽선 시계열이라 부른다.

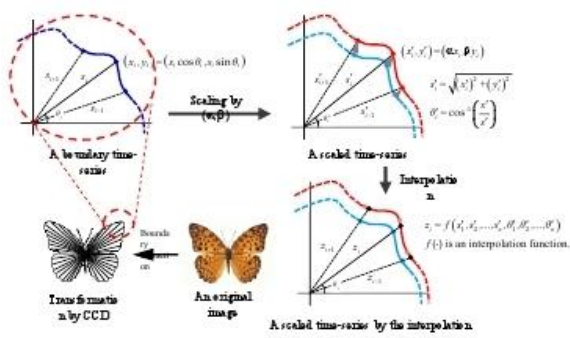
본 논문에서는 부분 노이즈 제거를 지원하는 색인 기반 윤곽선 매칭 시스템의 설계 및 구현을 다룬다. 이전 연구[4]에서는 윤곽선 이미지 내에 포함되어 있는 제한된 노이즈 즉, 부분 노이즈를 고려하는 윤곽선 매칭을 제안하였다. 이러한 부분 노이즈는 윤곽선 이미지 내에 크기, 길이, 위치에 따라 다양하게 존재한다. 질의 이미지인 경우, 부분 노이즈 제거는 간단하게 전처리과정에서 한 번만 실행하면 된다. 그러나, 데이터 이미지인 경우, 데이터 이미지들로부터 가능한 부분 노이즈들을 모두 고려해야 되기 때문에 매우 어려운 문제이다. 이 문제를 해결하기 위해서 이전 연구에서는 부분 노이즈 제거 윤곽선 매칭을 제안하였다. 그러나, 이 매칭 방법은 적어도 한 번은 모든 데이터 이미지에 접근해야 하기 때문에 디스크 I/O 오버헤드가 발생한다.

• First Author: Bum-Soo Kim, Corresponding Author: Bum-Soo Kim

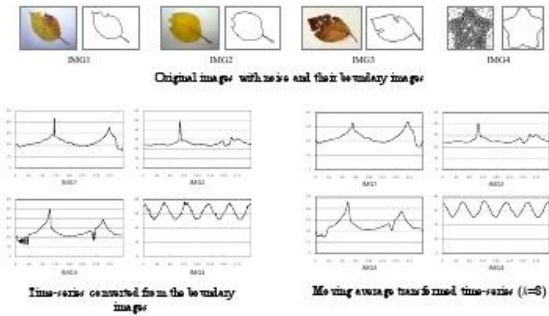
*Bum-Soo Kim (bumsookim@kict.re.kr) Dept. of Future Technology and Convergence Research, Korea Institute of Civil Engineering and Building Technology

*Jin-Uk Kim (jukim@kict.re.kr), Dept. of Future Technology and Convergence Research, Korea Institute of Civil Engineering and Building Technology

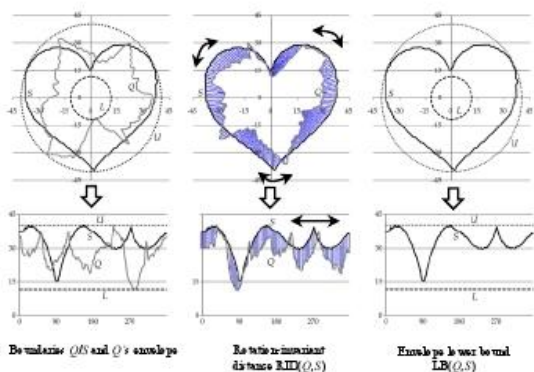
• Received: 2019. 04. 02, Revised: 2019. 04. 25, Accepted: 2019. 04. 25.



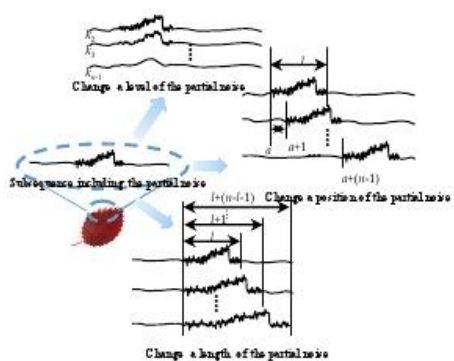
(a) Scaling a boundary image in scale-invariant boundary matching [1].



(b) Removing a whole noise from a boundary image in noise-control boundary matching [2].



(c) Rotating a boundary image in rotation-invariant boundary matching [3].



(d) Removing a partial noise from a boundary image in partial denoising boundary matching [4].

Fig. 1. Examples of various studies on boundary matching.

부분 노이즈 제거 윤곽선 매칭의 디스크 I/O 오버헤드 문제를 해결하기 위해 본 논문에서는 다차원 색인을 사용하여 색인 기반 접근법을 제안한다. 색인을 구축하기 위해, 각각 데이터 시계열마다 부분 노이즈 제거 시계열을 저장원 포인트로 변환한 후 본 논문에서는 이들 포인트들을 포함하는 저장원 MBR(minimum bounding rectangle)을 구성한다. 이때, R*-tree [5]와 같은 다차원 색인에서 구성된 MBR을 저장한다. 부분 노이즈 제거 시계열들 사이의 엔트리들은 서로 매우 유사하기 때문에 부분 노이즈 제거 시계열들을 MBR로 구성 가능하다. 즉, 부분 노이즈 제거 위치가 비록 다를지라도 부분 노이즈 제거 시계열들에서 많은 원본 엔트리들은 다른 부분 노이즈 제거 시계열에서 자주 사용된다.

본 논문에서는 사용자들의 편의를 위해 GUI 환경에서 클라이언트-서버 시스템을 설계하고 개발한다. 먼저 시스템은 윤곽선 이미지들로부터 변환된 데이터 시계열을 가진 윤곽선 데이터베이스를 구성한다. 이것은 빠른 매칭을 위해 1차원 변환 방법과 다차원 색인 기법을 활용한다. 클라이언트 시스템은 사용자에 의해 주어지는 질의 이미지를 질의 시계열로 변환한다. 그때, 질의 시계열과 함께 유사 척도로서 허용치 ϵ 와 노이즈제거 레벨 d , 노이즈제거 길이 l 를 서버로 보내진다. 서버는 질의 시계열, 주어진 허용치 ϵ , 노이즈제거 파라미터 값을 사용하여 윤곽선 데이터베이스 검색에 의해 유사한 윤곽선 이미지들을 얻게 된다. 이때, 유사한 이미지들을 클라이언트 시스템에 보내게 된다.

실험들을 통해 본 논문에서는 기존 방법과 걸린 시간을 비교하여 부분 노이즈 제거 윤곽선 매칭에서 제안한 색인기반 시스템의 우수성을 보인다. 또한, 제안한 방법은 부분 노이즈제거를 지원하는 매우 정확하고 직관적인 매칭을 제공함을 보여준다. 이러한 실험결과들을 기반으로 제안한 색인 기반 부분노이즈 제거 윤곽선 매칭 방법은 매우 실용적이고 사용자 친화적이라 확신한다.

II. Related works

2.1 Time-Series Matching

시계열 매칭은 시계열 데이터베이스에서 주어진 시계열과 유사한 시계열을 찾는 문제이다 [1, 6, 7]. 시계열 X 와 Y 가 있을 때, 그 유사척도 거리 $D(X, Y)$ 가 사용자에서 주어진 허용치 ϵ 보다 작거나 같으면 X 와 Y 는 유사하다고 정의한다 [1, 2, 6, 7]. 본 논문에서는 가장 간단하면서도 널리 사용하는 유클리디안 거리 유사척도 모델을 사용하며, ϵ 기반의 범위 질의에 초점을 둔다 [1, 2, 3]. 시계열 매칭 방법들은 크게 두 가지로 분류할 수 있다 [7]. 먼저, 전체 매칭(whole matching)은 질의 시계열과 데이터 시계열의 길이가 같은 시계열 매칭이다. 반면에, 부분시퀀스 매칭(subsequence matching)은 질의 시계열과 데이터 시계열의 길이가 서로 다르다.

색인 기반 시계열 매칭은 역시 이들 두 가지 시계열 매칭 방법에 따라 그 접근법이 다르다. Agrawal et al.[6]가 제안한 전체 매칭은 데이터 시계열들을 저차원 포인트들로 변환하여 그 포인트들을 색인에 직접 저장하는 것이 아니라 저차원 포인트들의 MBR만을 저장한 후 질의 시계열과 유사한 데이터 시계열을 색인에서 찾아 서로 비교하는 방법이다. Faloutsos et al.[7]가 제안한 서브시퀀스 매칭은 전체 매칭을 일반화한 방법이다. 서브시퀀스 매칭은 우선 데이터 시계열들을 슬라이드 윈도우 형태로 나눈 후 저차원 변환된 포인트들로 MBR을 구성하여 색인에 저장한다. 그 후 색인에서 질의 시계열과 유사한 서브시퀀스들을 찾아낸다. 본 논문에서는 윤곽선 매칭을 전체 매칭의 응용으로 활용하는데 초점을 둔다.

2.2 Image Matching

내용 기반 이미지 검색(CBIR)로 알려진 이미지 매칭은 이미지의 특성들(features)을 사용하여 주어진 질의 이미지와 유사한 데이터 이미지들을 찾는 문제이다 [8, 9, 10]. 이미지 매칭의 응용분야들은 사용하는 특성들에 따라 다르다. 본 논문에서는 이미지 매칭을 위해 컬러 또는 질감 특성들이 비슷한 값들을 가질 때 유용하게 사용할 수 있는 모양(shape) 특성들에 초점을 둔다 [9]. 제안한 방법에서는 예지, 곡률, 면적과 같은 모양 특성들 대신에 시계열 매칭 기술들을 활용하기 위하여 1차원 모양으로 표현이 가능한 가장 단순한 방법인 CCD(centroid contour distance) 방법 [1, 2, 11]을 사용한다. 그림 2은 CCD 방법에 의해 이미지를 시계열로 변환하는 예제를 보인다.

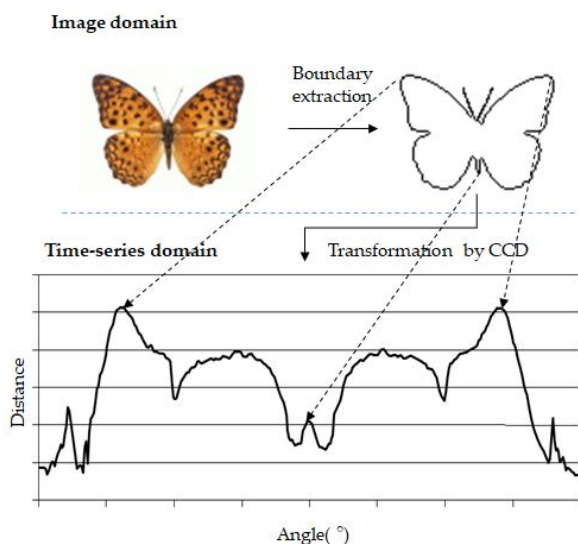


Fig. 2. An Example of converting an image to a time-series by CCD [4].

본 논문에서 다루는 노이즈 제거 문제는 Kim et al. [2, 4]에 의해 소개되었다. 이들은 먼저 대용량 이미지 데이터베이스에서 윤곽선 시계열(boundary time-series)에 있는 노이즈를 제거하는 것을 지원하는 윤곽선 매칭 즉, 전체 노이즈제거 윤곽선 매칭을 제안했다 [2]. 또한, 전체 노이즈제거에서 부분 노이즈

제거로 확장하여 지원하는 부분 노이즈제거 윤곽선 매칭을 다루었다 [4]. 전체 노이즈제거 윤곽선 매칭에서는 대용량 이미지 데이터베이스를 위해 색인 기반 매칭 방법을 활용하였다. 한편, 부분 노이즈제거 윤곽선 매칭에서는 시계열 매칭의 하한 기술을 활용한 효과적인 매칭 방법을 활용하였지만 대용량 이미지 데이터베이스를 위한 색인 방법은 고려되지 않았다. 즉, 이전 연구에서는 디스크 I/O 오버헤드 문제보다 매칭 기술에 초점을 두었다.

III. Design of an index-based system for partial denoising boundary matching

일반적으로 기존의 색인 방법들은 먼저 고차원 포인트들을 저차원 포인트들로 변환한 후 색인 안에 각각의 저차원 포인트들을 저장한다 [6, 12, 13]. 그러나, 이들 색인 방법들은 모든 부분 노이즈제거 시계열로부터 변환된 저차원 포인트들을 저장하는데 심각한 오버헤드가 발생하기 때문에 실용적이지 못하다. 대용량 이미지 데이터베이스에서 이 문제는 매우 심각하다. 결과적으로, 색인 기반 부분 노이즈제거 윤곽선 매칭을 위해 효율적인 색인 방법이 요구되어진다.

기본적인 색인 기반 매칭 시스템은 전처리, 범위 검색, 후처리 크게 세 가지 과정으로 이루어진다. 전처리 과정에서는 이동평균변환 [2]을 사용하여 데이터 시계열들의 부분 노이즈를 제거한다. 이 때, 데이터 시계열의 서브시퀀스들을 노이즈를 제거한 부분시퀀스로 대체된 시계열을 부분 노이즈제거 시계열이라 한다. 이러한 길이가 n 인 각각의 부분 노이즈제거 시계열을 f -차원 포인트로($f \ll n$) 저차원 변환한 다음, 변환된 포인트들은 f -차원의 R^* -트리 [2, 5]에 저장한다. 범위 검색 과정에서는 질의 시계열을 f -차원 포인트로 변환한 다음 허용치 ϵ 와 함께 범위 질의를 구성한다. 구성된 범위 질의는 색인에서 질의 시계열과 잠재적으로 유사한 후보 시계열(candidates)을 우선 판별한다. 여기서 후보 시계열이 발생하는 이유는 저차원 변환이 착오기각(false dismissal)은 발생하지 않음을 보장하지만 착오해답(false alarms)은 발생하기 때문이다 [2, 6, 7]. 따라서, 후처리 과정에서는 각각의 후보 시계열에서 유사한 데이터 시계열들을 찾아내기 위해 실제 디스크에 있는 데이터 시계열과 질의 시계열을 비교한다. 즉, 질의 시계열과 후보 시계열에 해당하는 실제 데이터 시계열과 유사거리를 계산하여 착오해답인 후보 시계열들을 제거해 나간다.

후처리 과정에서 발생하는 오버헤드 문제를 해결하기 위해서는 색인에서 f -차원 포인트들을 직접 저장하는 것이 아니라 f -차원 MBR을 저장한다 [2]. 이는 부분 노이즈제거 시계열에서 노이즈를 제거한 서브시퀀스를 제외한 나머지 엔트리들은 심지어 노이즈제거 위치가 다를지라도 다른 부분 노이즈제거 시계열에서 빈번히 사용된다. 즉, 하나의 원본 윤곽선 이미지로부터 생성되어 나온 부분 노이즈제거 시계열들은 각각 서로 유사하다. 따라서, 본 논문에서는 이러한 직관에 의해 각각의 부분 노이즈제거 시계열들을 색인에 직접 저장하는 대신에 하나의 윤곽선 시계열에서 나온 모든

부분 노이즈제거 시계열들 하나의 MBR로 구성하여 색인 검색이 가능하도록 하는 접근법을 제안한다.

그림 3은 부분 노이즈제거 윤곽선 매칭을 위한 색인 기반 시스템의 전체 아키텍처이다. 그림을 보면, 시스템은 클라이언트와 서버로 나뉜다. 클라이언트 쪽에서는 윤곽선 추출, 윤곽선 시계열 변환, 부분 노이즈제거 시계열 생성, 저차원 변환, MBR 구성, 색인 구성 등 모든 전처리 과정이 처리된다. 여기에서 전처리 과정들은 오직 한번만 수행된다. 반면에, 서버 쪽에서는 범위 질의를 통한 범위 검색을 수행하여 후처리 작업 통한 매칭 결과를 얻어내어 클라이언트에 전달하는 역할을 한다. 윤곽선 데이터베이스는 데이터 이미지들의 윤곽선 시계열들의 저장소를 의미하며 후처리 과정에서 후보 시계열들의 실제 윤곽선 시계열들과 비교할 때 사용된다. 제안한 시스템은 색인을 통해 실시간 매칭 결과를 얻을 수 있어 클라이언트-서버 형태로 다수의 사용자들에게 서비스를 제공할 수 있다.

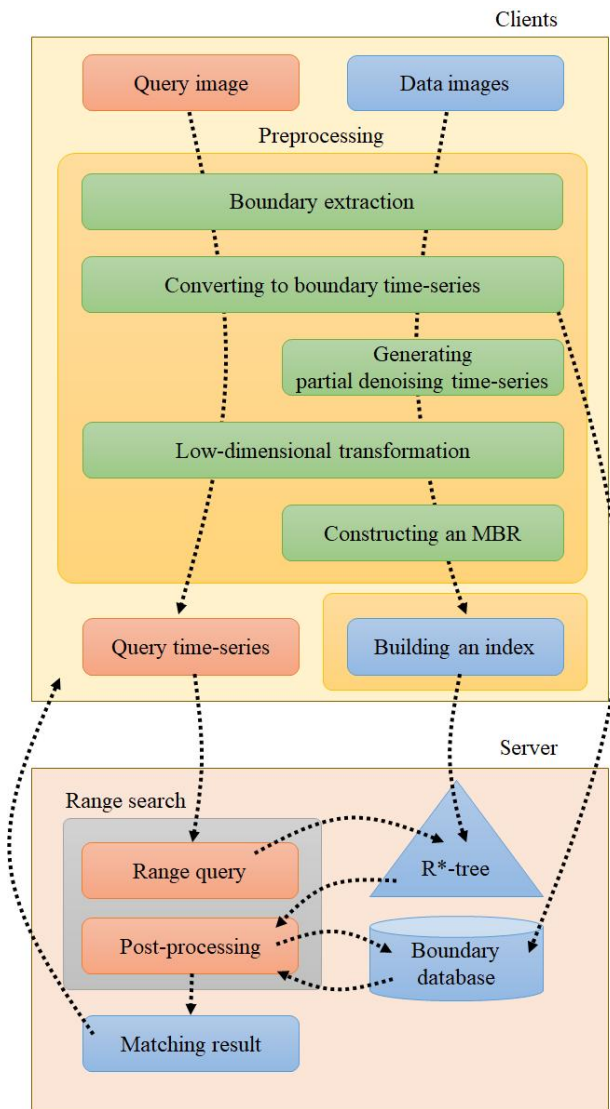
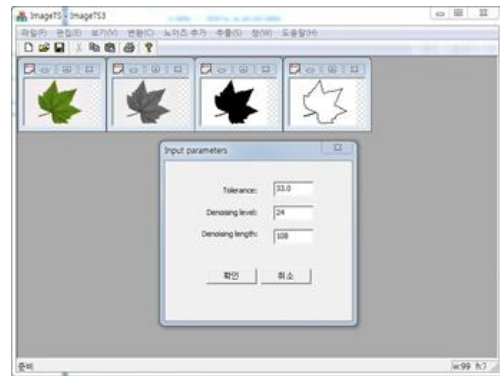


Fig. 3. An overall architecture of an index-based system for partial denoising boundary matching.

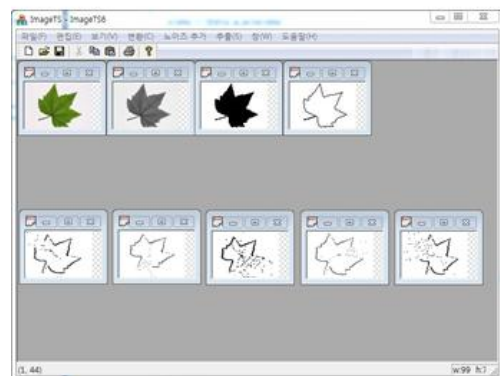
IV. Implementation of an index-based system for partial denoising boundary matching

제안한 시스템에서 이미지에서 윤곽선 시계열을 변환시키는 과정은 (1) 그레이스케일 이미지의 바이너리 변환, (2) 바이너리 이미지의 윤곽선 탐색 등 크게 두 가지로 나눌 수 있다. 바이너리 변환 과정에서는 먼저 칼라 이미지를 그레이스케일 이미지로 변환 후 그레이스케일 이미지로부터 바이너리 이미지를 얻는다. 이 과정에서는 바이너리 변환의 정도를 조절할 수 있는 바이너리 한계치가 필요하다. 실험에서는 바이너리 한계치로써 240을 사용하였다. 바이너리 이미지에서 윤곽선을 탐색하는 과정에서 본 논문에서는 8-이웃 연결성(8-neighborhood connectivity) 알고리즘을 사용하였다 [10].

본 논문에서는 색인 기반의 부분 노이즈제거 윤곽선 매칭을 위한 클라이언트-서버 시스템을 구현하였다. 그림 4는 제안한 시스템의 입력 화면과 출력 화면을 나타낸다. 그림 4(a)와 같이 부분 노이즈제거 윤곽선 매칭의 결과를 얻기 위해서는 허용치와 노이즈제거 파라미터 값들을 질의 이미지와 함께 서버 쪽으로 보내야 한다. 그림 4(b)에서는 주어진 질의 이미지와 유사한 결과 이미지들을 보인다.



(a) An input screenshot of the matching system.



(b) An output screenshot of the matching system.

Fig. 4. Example screenshots of the proposed boundary matching system.

성능실험에서는 이전 연구 [4]에서 사용한 실험 윤곽선 데이터셋을 사용한다. 이 데이터셋은 각 원본 이미지로부터 길이나 위치를 달리하여 아홉 가지 경우의 부분 노이즈들을 포함하는 길이 360의 총 102,590 개의 윤곽선 시계열들이다. 부분 노이즈를 생성하기 위해서는 가우시안 모델 [1, 2, 3, 4]을 사용한다. 이 원본 이미지들은 모두 인터넷으로부터 수집된 오브젝트 이미지들이다. 이 데이터셋은 비록 약 1,000개의 원본 이미지 [1, 2, 3, 4]만을 사용하지만, 하나의 원본 이미지에 여러 개의 윤곽선 이미지들이 포함되어 있을 수 있기 때문에 CCD 방법에 의해 생성된 윤곽선 시계열들은 사실상 더 많게 된다.

성능실험은 서버 쪽에서 수행하였다. 성능실험을 위해 본 논문에서는 2.0GHz Intel Core 2 Duo CPU, 2.0GB 램, 500GB 하드 디스크 사양의 PC를 사용하였다. 운영체제는 CentOS 6.3 이며, C/C++ 언어로 색인 구축 및 매칭 알고리즘을 개발하였다. 다차원 색인으로는 R*-tree를 사용하였으며, 색인 내의 색인 및 데이터 페이지 사이즈는 4,096 바이트로 설정하였다. 또한, 저장원 변환 알고리즘은 PAA [1]를 사용하였으며, 이를 가지고 시계열을 72 특성들로 변환하였다.

본 논문에서는 부분 노이즈제거 윤곽선 매칭에서 기본 매칭 알고리즘(NIV-OG)[4]과 제안한 색인 기반 매칭 알고리즘(BBI)의 걸린 시간을 서로 비교하였다. 그림 5는 매칭 알고리즘들의 확장성을 보인다. (Y축은 로그스케일이다.) 그림을 보면, 색인 기반 매칭 알고리즘은 선형적으로 확장됨을 알 수 있다. 즉, 데이터 시계열들의 수가 증가할수록, 색인 기반 매칭 알고리즘은 대용량 윤곽선 데이터베이스를 다루는데 적합하다는 것을 의미한다. 따라서, 제안한 시스템은 대용량 윤곽선 데이터베이스일지라도 수 초안에 결과를 얻을 수 있다. 그림 5의 실험 결과를 보면, 색인 기반 매칭 알고리즘은 기본 매칭 알고리즘에 비해 약 2580.8 배 매칭수행 성능이 우수함을 보인다. 결과적으로, 본 논문에서 제안한 색인 기반 부분 노이즈제거 윤곽선 매칭 시스템은 대용량 윤곽선 데이터베이스를 다루는데 매우 효과적인 방법이라 사료된다.

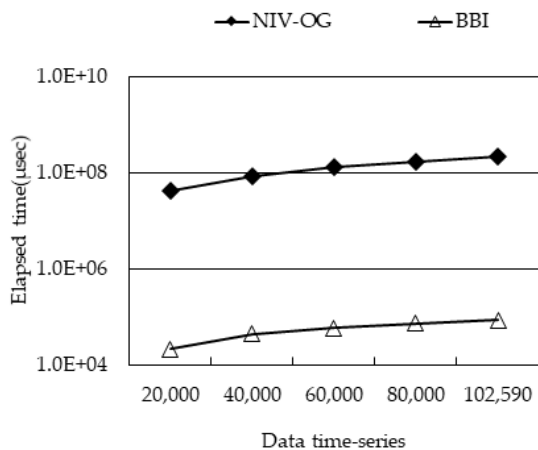


Fig. 5. The scalability of the naive matching and the index-based matching algorithms.

V. Conclusions

본 논문에서는 색인 기반의 부분 노이즈제거 윤곽선 매칭 시스템을 설계 및 구현하였다. 다차원 색인에서 윤곽선 시계열들을 저장하기 위한 수행 오버헤드 문제를 해결하기 위해, 제안한 시스템에서는 시계열 매칭 기술을 활용하였다. 제안한 시스템은 실시간으로 대용량 데이터베이스를 다수의 사용자가 쉽게 사용할 수 있도록 클라이언트-서버 모델 기반으로 구현하였다. 또한, 제안한 색인 기반 매칭 알고리즘은 부분 노이즈제거 윤곽선 매칭의 기본 매칭 알고리즘보다 우수한 성능을 보임을 확인하였다. 본 논문에서는 매칭 결과들을 식별하기 쉽고 직관적으로 이해하는 GUI 기능 개발에 초점을 두었다. 향후 연구로는 범위 질의 대신에 k-NN 질의를 지원하는 시스템 설계와 구현을 할 예정이다.

REFERENCES

- [1] Y.-S. Moon, B.-S. Kim, M. S. Kim, and K.-Y. Whang, "Scaling-Invariant Boundary Image Matching Using Time-Series Matching Techniques," *Data & Knowledge Engineering*, Vol. 69, No. 10, pp. 1022-1042, Oct. 2010.
- [2] B.-S. Kim, Y.-S. Moon, M.-J. Choi, and J. Kim, "Interactive Noise-Controlled Boundary Image Matching Using the Time-Series Moving Average Transform," *Multimedia Tools and Applications*, Vol. 72, No. 3, pp. 2543-2571, Oct. 2014.
- [3] W.-K. Loh, S.-P. Kim, S.-K. Hong, and Y.-S. Moon, "Envelope-based Boundary Image Matching for Smart Devices under Arbitrary Rotations," *Multimedia Systems*, Vol. 21, No. 1, pp. 29-47, Feb. 2015.
- [4] B.-S. Kim, Y.-S. Moon, and J.-G. Lee, "Boundary Image Matching Supporting Partial Denoising Using Time-Series Matching Techniques," *Multimedia Tools and Applications*, Vol. 76, No. 6, pp. 8471-8496, Mar. 2017.
- [5] N. Beckmann, H.-P. Kriegel, R. Schneider, and B. Seeger, "The R*-tree: An Efficient and Robust Access Method for Points and Rectangles," In *Proc. of the ACM SIGMOD Int'l Conf. on Management of Data*, pp. 322-331, Atlantic City, New Jersey, May 1990.
- [6] R. Agrawal, C. Faloutsos, and A. Swami, "Efficient Similarity Search in Sequence Databases," In *Proc. the 4th Int'l Conf. on Foundations of Data Organization and Algorithms*, pp. 69-84, Chicago, Illinois, Oct. 1993.
- [7] C. Faloutsos, M. Ranganathan, and Y. Manolopoulos, "Fast Subsequence Matching in Time-Series Databases," In *Proc. of the ACM SIGMOD Int'l Conf. on Management of Data*, pp. 419-429, Minneapolis, Minnesota, May 1994.

- [8] S. Belongie, J. Malik, J. Puzicha "Shape Matching and Object Recognition Using Shape Contexts," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 24, No. 4, pp. 509-522, Apr. 2002.
- [9] R. Datta, D. Joshi, J. Li, and J. Z. Wang, "Image Retrieval: Ideas, Influences, and Trends of the New Age," *ACM Computing Surveys*, Vol. 40, No. 2, pp. 34-94, Apr. 2008.
- [10] M. Sonka, V. Hlavac, and R. Boyle, *Image Processing, Analysis, and Machine Vision*, 4th Ed., Cengage Learning, 2014.
- [11] D. Zhang and G. Lu, "Review of Shape Representation and Description Techniques," *Pattern Recognition*, Vol. 37, No. 1, pp. 1-19, Jan. 2004.
- [12] S. Berchtold, C. Bohm, and H.-P. Kriegel, "The Pyramid-Technique: Towards Breaking the Curse of Dimensionality," In *Proc. of the ACM SIGMOD Int'l Conf. on Management of Data*, pp. 142-153, Seattle, Washington, June 1998.
- [13] R. Arandjelović and A. Zisserman, "Three Things Everyone Should Know to Improve Object Retrieval," In *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 2911-2918, Providence, Rhode Island, June 2012.

Authors



Bum-Soo Kim received his Ph. D. (2013) degrees in computer science from Kangwon National University. From 2013 to 2017, he was a postdoctoral researcher in Korea Advanced Institute of Science and Technology (2013 and 2015), Kangwon

National University (2014), and Korea University (2016-2017). He is currently a postdoctoral researcher in Department of Future Technology and Convergence Research from Korea Institute of Civil Engineering and Building Technology (KICT). His research interests include time-series data mining, construction bigdata analysis, and data mining applications.



Jin-Uk Kim received the B.S., M.S. and complete Ph.D. in Computer Science and Engineering from ChungNam National University, Korea, in 1988, 1991 and 1999, respectively Kim is currently a Research fellow in the Korea Institute of Civil

Engineering and Building Technology (KICT). He is interested in Database, Bigdata, Information Retrieval.