

CSI: PC² - A High Performance Biometric System

Samir Brulic, Samira Brulic, Pascal Deppe, Dominic Eschweiler, Rosbeh Etemadi,
Klaus Herbold, Christoph Konersmann, Robert Meiche, Christoph Scholz,
Nils Timm, André Ückermann, Elmar Weber, René C. Zorn

{samir | sbrulic | bluesky | weiler | frost77 | klausher | c.k | rmeiche |
cscholz | timm84 | auecker | eweber | zorn }@uni-paderborn.de

1 Einführung

Biometrische Systeme gewinnen, vor allem im Bereich der Kriminalitätsverfolgung und Sicherheit, immer mehr an Bedeutung. Mit zunehmender Verbreitung steigen auch die Anforderungen an biometrische Systeme: Die Anzahl der biometrischen Templates erreicht zig Millionen und Anfragen müssen in sekundenschnelle beantwortet werden. Aktuelle Systeme sind hierfür nicht ausgelegt. So braucht z. B. das System des FBI's zwischen 2 (Polizeibehörden) und 24 (zivile Behörde) Stunden um eine Anfrage auf 50.000.000 Fingerabdrücke zu beantworten[FBoIF05].

In der Projektgruppe CSI: PC²[CSI07] entwickeln seit April 2007 13 Studenten ein biometrisches System, welches durch Cluster Parallelisierung und unter Verwendung von FPGAs dieses Problem angeht. In dem Projekt werden Fingerabdrücke und Iriden als biometrische Merkmale betrachtet.

Die Ziele der Projektgruppe sind zwei: Hohe Performance und gute Erkennungsgenauigkeit. Eine hohe Performance ist nötig bei der Identifikation einer Person anhand von biometrischen Merkmalen, d. h. wenn z. B. ein am Tatort gefundener Fingerabdruck gegen eine Datenbank mit mehreren Millionen Abdrücken verglichen wird. Eine hohe Erkennungsgenauigkeit wiederum ist speziell dann erforderlich, wenn es um die Verifikation einer Person geht, dazu müssen vor allem „false positives“ verhindert werden.

2 Realisierung

Das komplette System wurde in C/C++ realisiert. Abbildung 1 zeigt eine grafische Darstellung der Architektur. Auf der Client Seite existieren Module um Scanner für Fingerabdrücke und Iriden anzusprechen, eine GUI zur Steuerung sowie ein Controller Modul welches den Ablauf koordiniert und die Anfragen an den Server übermittelt. Die Anfragen werden verschlüsselt an den Server gesendet wo der Master Cluster Control Daemon (CCD) die Anfrage entgegen nimmt und per MPI an alle verfügbaren Knoten verteilt.

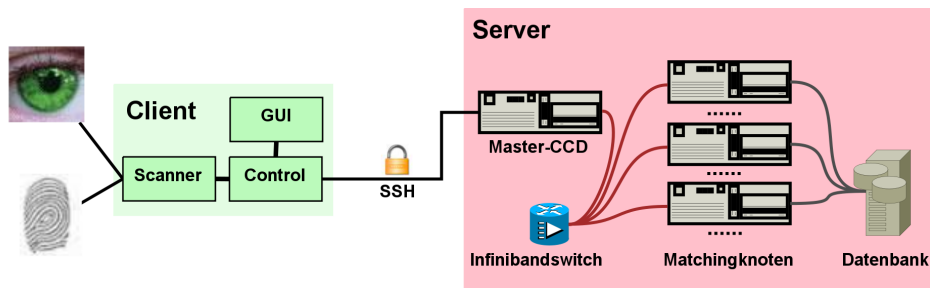


Abbildung 1: Übersicht über die gesamte Architektur des Systems

Anschließend vergleicht jeder Knoten das Template der Anfrage gegen seinen Teil der Templatedatenbank. Alle Ergebnisse über einem gewissen Schwellwert werden per MPI zurück an den Master-CCD geschickt, welcher die besten kombiniert und das Ergebnis an den Client zurückschickt.

Auf der Cluster Seite wird beim Start die komplette Datenbasis gelesen und per MPI im Round Robin Stil an alle Knoten verteilt. Da jeder Knoten nur einen Teil der Templatedatenbank erhält können die Templates i. d. R. problemlos im RAM gehalten werden.

Neben der Möglichkeit Anfragen zu stellen kann der Client auch die Datenbasis des Servers um neue Templates erweitern. Hierzu wird ein vom Scanner aufgenommenes Bild des biometrischen Merkmals (Fingerabdruck oder Iris) an den CCD weitergeleitet. Der CCD führt dann eine Feature Extraction durch, wandelt also das Bild in ein Template für den späteren Vergleich um und speichert das Template zusammen mit dem Bild in der Datenbasis.

Die Datenbasis besteht aus den aufgenommenen Bildern und den extrahierten Templates für Fingerabdrücke und Iriden. Die Templates werden binär kodiert hintereinander in eine Flatfile Datenbank geschrieben um ein effizientes Auslesen der Datenbank beim Start des Servers zu ermöglichen. Um die Startzeit des Servers weiterhin zu optimieren besteht die Möglichkeit die Datenbank auf mehrere Dateien zu verteilen, so dass ein paralleles Einlesen möglich ist.

Durch die Verwendung einer Client/Server Architektur ist es problemlos möglich den Client auszutauschen um z. B. sowohl ein Terminal als auch ein mobiles Gerät mit dem gleichen Server zu nutzen.

Wie anfangs erwähnt wollen wir eine Beschleunigung des Matchingvorgangs erreichen indem wir zusätzlich FPGAs verwenden. Hierzu sind einige Knoten des PC² Clusters mit FPGA Boards von Alpha Data [Dat08] ausgestattet welche über den PCI-X Bus mit dem System kommunizieren. Die Beschleunigung der FPGA Variante der Matchingalgorithmen gegenüber der Software Variante basiert im Prinzip darauf, dass wir auf einem FPGA die Möglichkeit haben mehrere Attribute parallel zu vergleichen. Hierdurch ist ein hoher Speedup gegenüber der Software Variante möglich (siehe Abschnitt 3).

Neben der Architektur sind die verwendeten Matchingalgorithmen mit der wichtigste Fak-

tor für die Performance und vor allem für die Erkennungsgenauigkeit. Für die Fingerabdrücke verfolgen wir drei verschiedene Ansätze. Der erste ist der in der Praxis weitverbreitete und erprobte Ansatz der die besonderen Punkte eines Fingerabdrucks (Minutien) miteinander vergleicht[RJR95, JFRQ06]. Als zweiten Ansatz setzen wir auf ein Clustering Verfahren, welches durch eine Baumstruktur die Suchmenge in wenigen Schritten von mehreren Millionen auf einige Tausend reduziert[LJK07]. Der dritte Ansatz ist ein relativ neuer Matching Ansatz welcher auf einem Bildfilter basiert, der die Grauwerte der Fingerabdrücke vergleicht[MJ02]. Durch die Verwendung von drei Ansätzen bauen wir auf eine Erhöhung der Erkennungsgenauigkeit durch einen Kreuzvergleich aller Verfahren. Weiterhin haben wir eine Auswahl und können den besten Algorithmus nutzen. Für das Matching der Iriden verwenden wir den Algorithmus von Daugman, dem de facto Standard in dem Bereich[Dau04].

3 Bisherige Ergebnisse und Ausblick

Die bisherigen Ergebnisse unserer Projektgruppe sind viel versprechend. Auf unserem Testcluster¹ erreichen wir pro CPU über 60.000 Vergleiche pro Sekunde bei den Fingerabdrücken. Konkret bedeutet dies, dass die unoptimierte Software Variante des auf Minutien basierenden Algorithmus (Minutenmatcher) es schafft, eine Datenbasis von 50.000.000 Fingerabdrücken auf 32 Knoten (64 CPUs) innerhalb von 13 Sekunden zu vergleichen. Die FPGA Variante des gleichen Algorithmus beschleunigt das ganze um den Faktor 27.

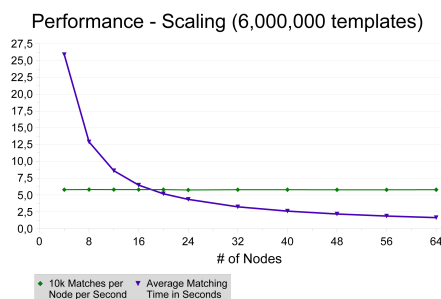


Abbildung 2: Skalierung und Durchsatz

Viel versprechende Ergebnisse erhalten wir auch bei dem Clustering Verfahren für Fingerabdrücke. Auf zwei Knoten (4 CPUs) und einer Datenbank mit 2.000.000 Fingerabdrücken dauert ein Vergleich nur 0,25 Sekunden. Da das Clustering unter Umständen mehrere Tausend Fingerabdrücke zurückliefert werden diese anschließend mit dem Minutenmatcher verglichen um eine bessere Erkennungsgenauigkeit zu erreichen.

Ein weiteres wichtiges Ergebnis ist, dass sich unser System proportional zur Anzahl der Clusterknoten skaliert, man also beliebig viele Knoten nehmen kann ohne Performance zu verlieren (vgl. Abbildung 2).

Der Abschluss der Projektgruppe CSI: PC² ist Mitte April 2008, bis dahin werden wir weiter an der Erkennungsgenauigkeit und der Performance arbeiten. Aktuelle Ergebnisse und Entwicklungen sowie weitere Dokumentation sind auf der Homepage der Projektgruppe CSI:PC²[CSI07] zu finden.

¹Arminius Cluster des PC², 200 Knoten mit jeweils zwei 64bit Xeon CPUs mit 3.2 GHz, weitere Details siehe [fPCP08]

Literatur

- [CSI07] Homepage Projektgruppe CSI:PC², 2007. <http://wwwcs.uni-paderborn.de/pc2/teaching/pg-csipc2.html>.
- [Dat08] Alpha Data, 2008, (accessed February 16, 2008). <http://www.alpha-data.com/>.
- [Dau04] John Daugman. How Iris Recognition Works. In *IEEE Transaction on Circuits and Systems for Video Technology*, Jgg. 14 (1), Seiten 21–30, 2004.
- [FBoIF05] Criminal Justice Information Services (CJIS) Division Federal Bureau of Investigation (FBI). Integrated Automated Fingerprint Identification System or IAFIS, 2005. <http://www.fbi.gov/hq/cjisd/iafis.htm>.
- [fPCP08] Paderborn Center for Parallel Computing (PC²). Architecture of the ARMINIUS cluster, 2006, (accessed February 16, 2008). <http://wwwcs.uni-paderborn.de/pc2/services/systems/arminius/system-architecture.html>.
- [JFRQ06] Ying Jie, Yuan Yi Fang, Zhang Renjie und Song Qifa. Rapid and brief communication: Fingerprint minutiae matching algorithm for real time system. *Pattern Recogn.*, 39(1):143–146, 2006.
- [LJK07] Manhua Liu, Xudong Jiang und Alex Chichung Kot. Efficient fingerprint search based on database clustering. *Pattern Recogn.*, 40(6):1793–1803, 2007.
- [MJ02] Muhammad Umar Munir und Dr. Muhammad Younus Javed. Fingerprint Matching Using Gabor Filters. *National Conference on Emerging Technologies NCET*, 2002.
- [RJR95] Nalini K. Ratha, Anil K. Jain und Diane T. Rover. An FPGA-based point pattern matching processor with application to fingerprint matching. In *CAMP '95: Proceedings of the Computer Architectures for Machine Perception*, Seite 394, Washington, DC, USA, 1995. IEEE Computer Society.