

# Ovládanie počítačových aplikácií pomocou gest ruky



Michal Hozza



25.06.2012

Vedúci: RNDr. Marek Nagy

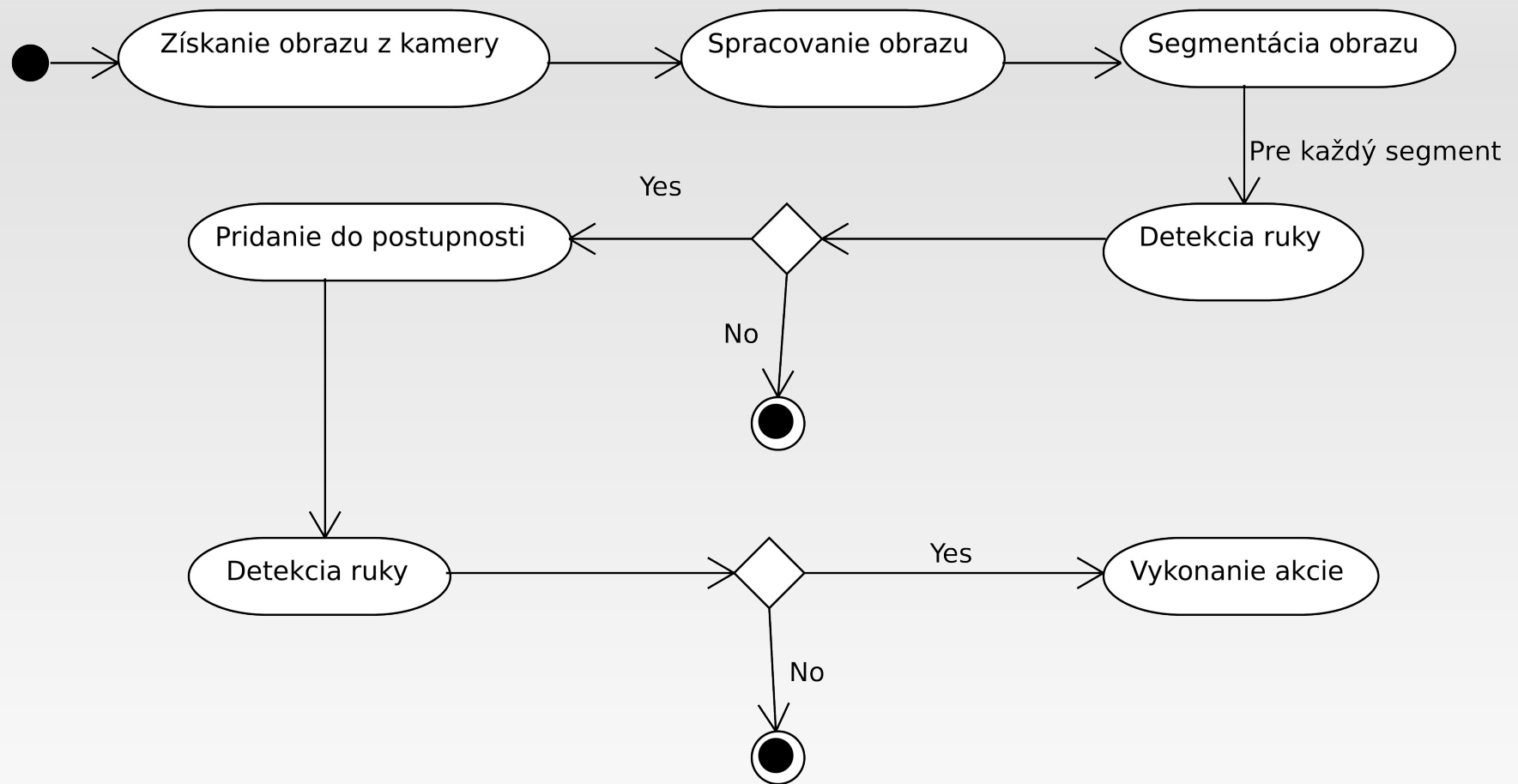
# Cieľ

- Pomocou počítačových neurónových sietí rozpoznávať jednoduché dynamické gestá prezentované rukou a zaznamenané webovou kamerou. Porovnať dva prístupy k rozpoznávaniu: cez doprednú a rekurentnú neurónovú sieť.
- Realizovaný rozpoznávač prepojiť na ovládanie aplikácií.

# Použitý jazyk a knihovny

- Linux
- C++
- Qt framework
- V4l2, Xtst, fftw3

# Základný algoritmus



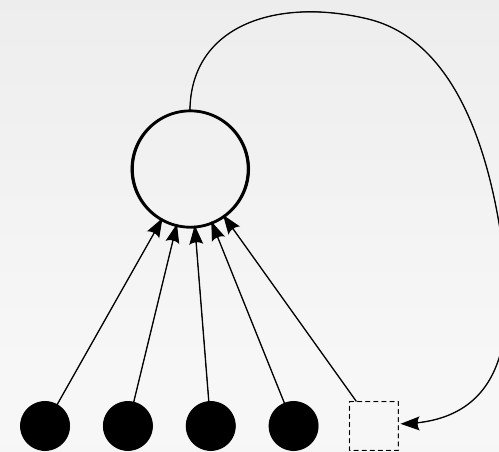
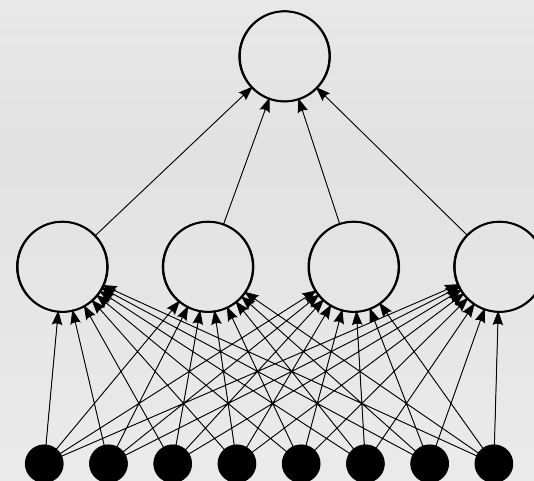
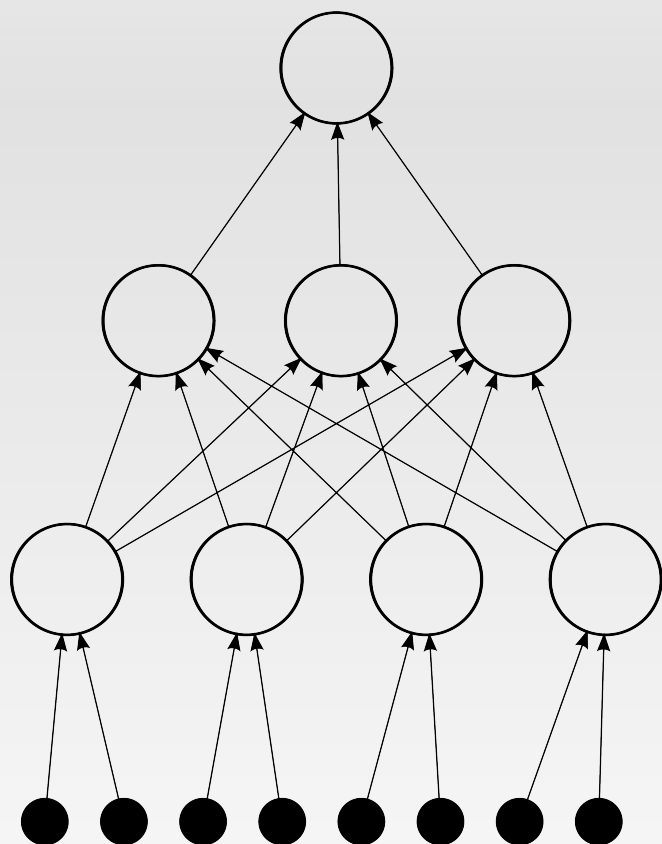
# Predspracovanie obrazu

- Rozdiel 2 po sebe idúcich obrázkov
- Vyčistenie
  - Príliš malé rozdiely zanedbáme
- Segmentácia
  - Rozpitie a následné použitie BFS
- Škálovanie (bilineárna interpolácia)
  - Neurónová sieť potrebuje presnú veľkosť vstupu
- Fourierova transformácia

# Detekcia ruky

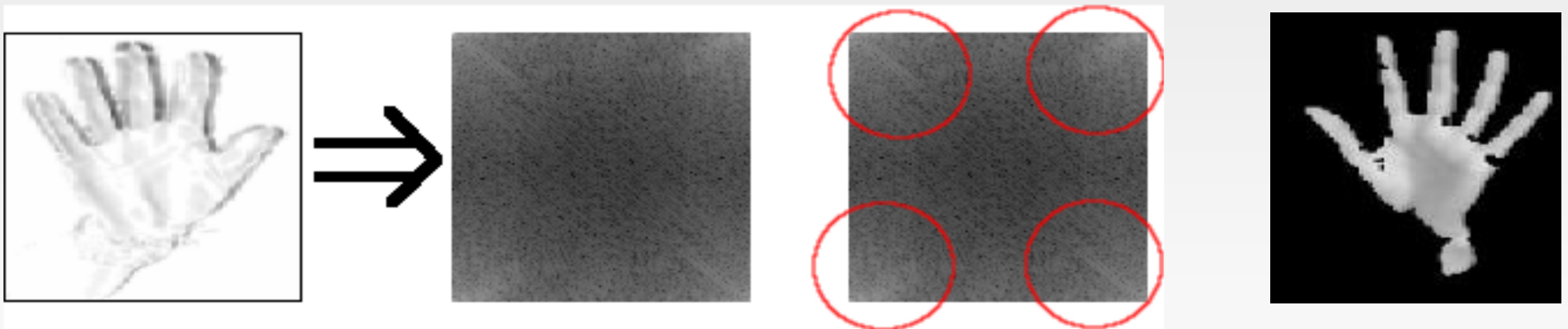
- Použitie neurónovej siete
- Rôzne typy
  - Dopredná neurónová sieť - rôzne architektúry
    - Spodná vrstva dostáva celý obrázok
    - Rozdelenie na časti a pridelenie vstupných neurónov k jednotlivým častiam
  - Použitie rekurentných neurónov

# Typy architektúr



# Detekcia ruky 2

- Rôzne dáta
  - Fourierova transformácia
  - Rozdielový obrázok
  - Obrázok z webkamery vybratý pomocou floodfill selekcie





# Detekcia gesta

- Postupnosť bodov
- 4 typy gest
  - Testuje sa každé gesto
- Jednoduchý algoritmus
  - Testovanie rozmerov opísaného obdĺžnika

# Porovnanie prístupov

## Porovnanie architektúr

Typ	Počet neurónov	Úspešnosť
Typ 1	47	65,07%
Typ 2	176(11)	67,47%
Typ 2	7;176(11)	67,37%
Typ 3	176(11)	<b>67,88%</b>

### Legenda:

- *Typ 1* = Viacvrstvová dopredná neurónová sieť
- *Typ 2* = Upravená verzia viacvrstvovej doprednej neurónovej siete
- *Typ 3* = Rekurentná neurónová sieť

## Vplyv Fourierovej transformácie

Typ dát	Úspešnosť
Obrázok (rozdielový)	65,06%
Fourierova transformácia	<b>67,37%</b>

## Rozdielový vs. pôvodný obraz

Typ dát	Úspešnosť
Rozdielový obraz	76,64%
Pôvodný obraz	66,51%
Rozdielový obraz - FT	<b>81,78%</b>
Pôvodný obraz - FT	74,73%

# Implementácia

- Knižnica pre neurónové siete
- Triedy na obrázky
- Trieda na spracovanie obrazu
- Trieda na rozpoznanie ruky
- Trieda na rozpoznanie gesta
- Pomocné programy

# Záver

- Aplikácia s úspešnosťou rozpoznávania 4 základných gest približne **80%**
- Ukázané metódy predspracovania obrazu a zhodnotený ich vplyv
- Ukázaný návrh architektúr neurónových sietí a porovnanie navrhnutých architektúr
- Možnosti do budúcnosti:
  - *Kalmanov filter*
  - Použitie *Skrytých Markvovych Modelov*
    - Namiesto (rekurentných) neurónových sietí
    - Na rozpoznanie gesta

# Ďakujem za pozornosť

Otázky?

Adresa projektu: <https://github.com/mhozza/HandControl>