

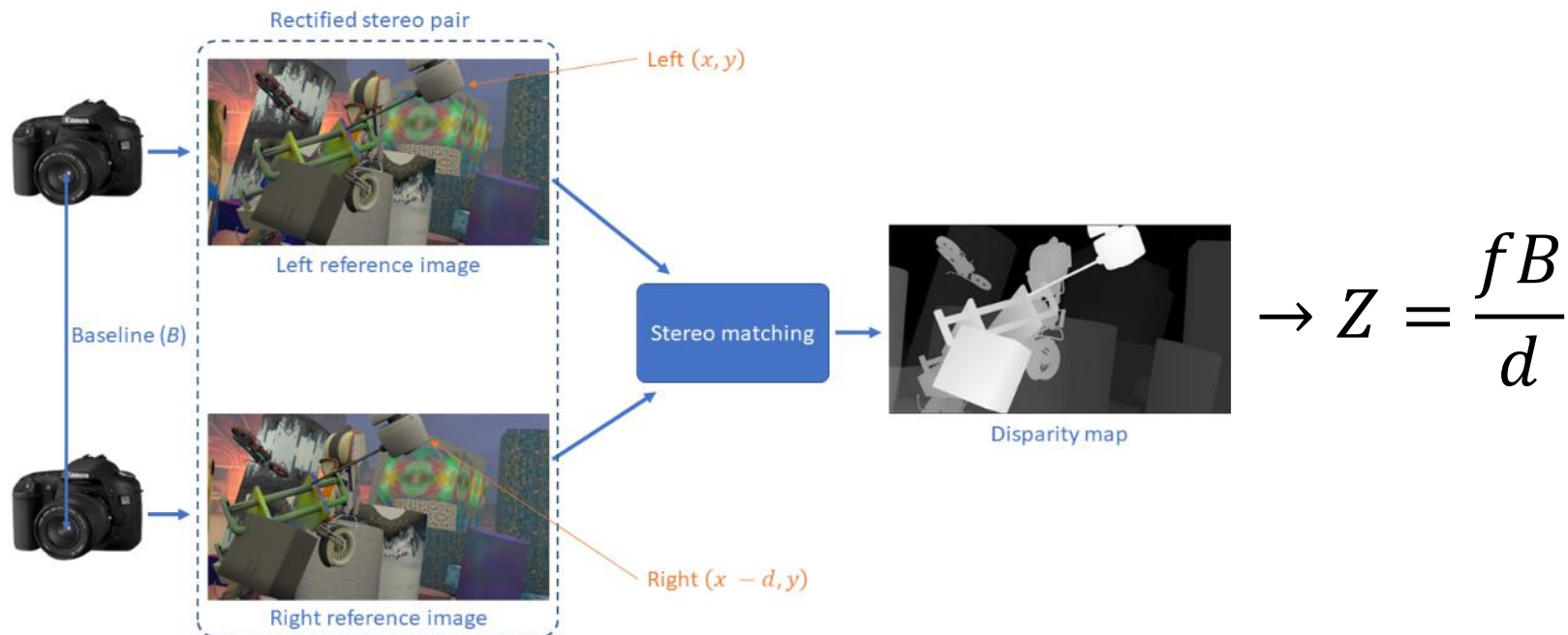
# Одређивање дубине из једне референтне слике оптимизовано за графичке процесоре

Студент: Душан Ердељан SW-43/2018

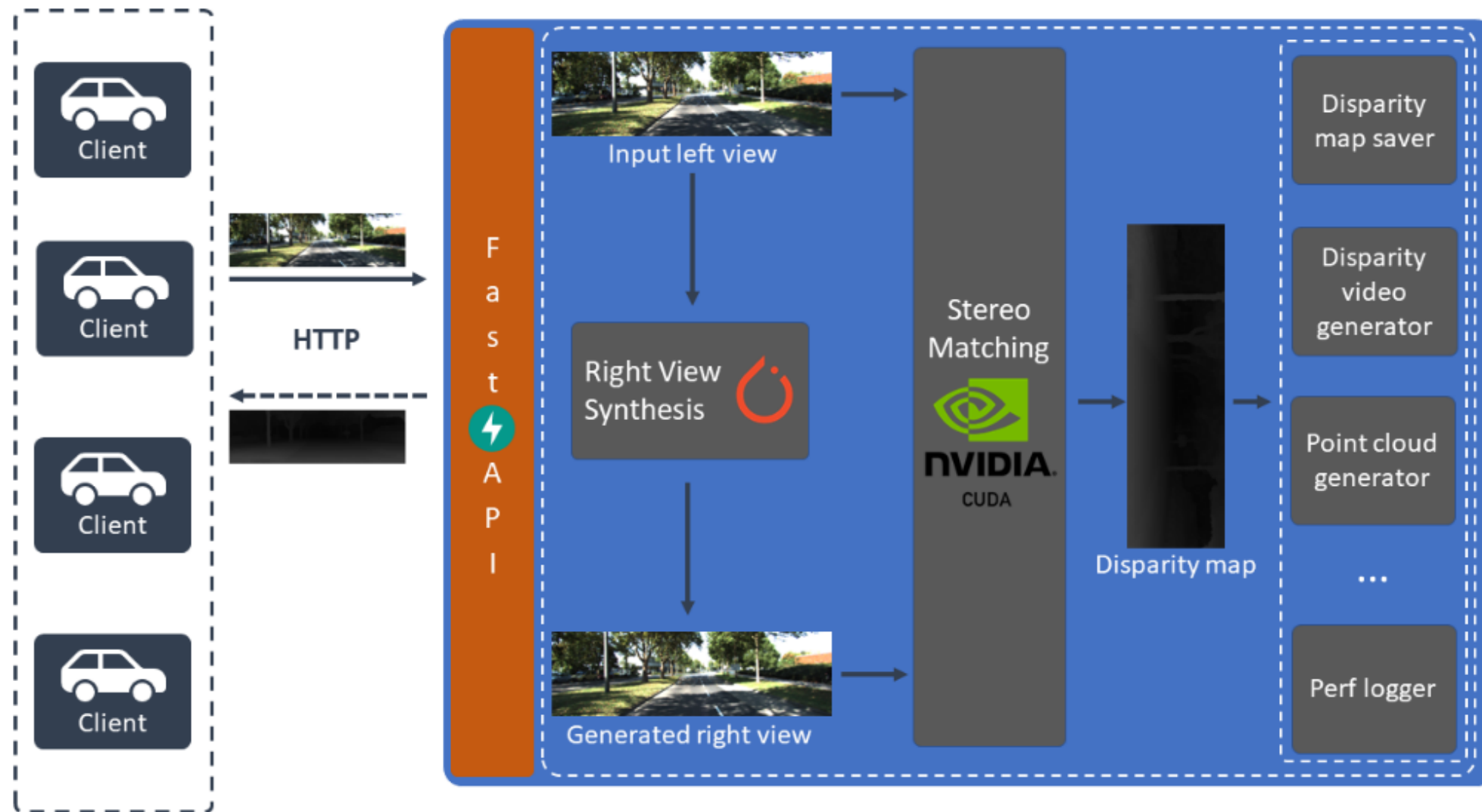
Ментор: др Јелена Сливка

# Мотивација и стерео визија

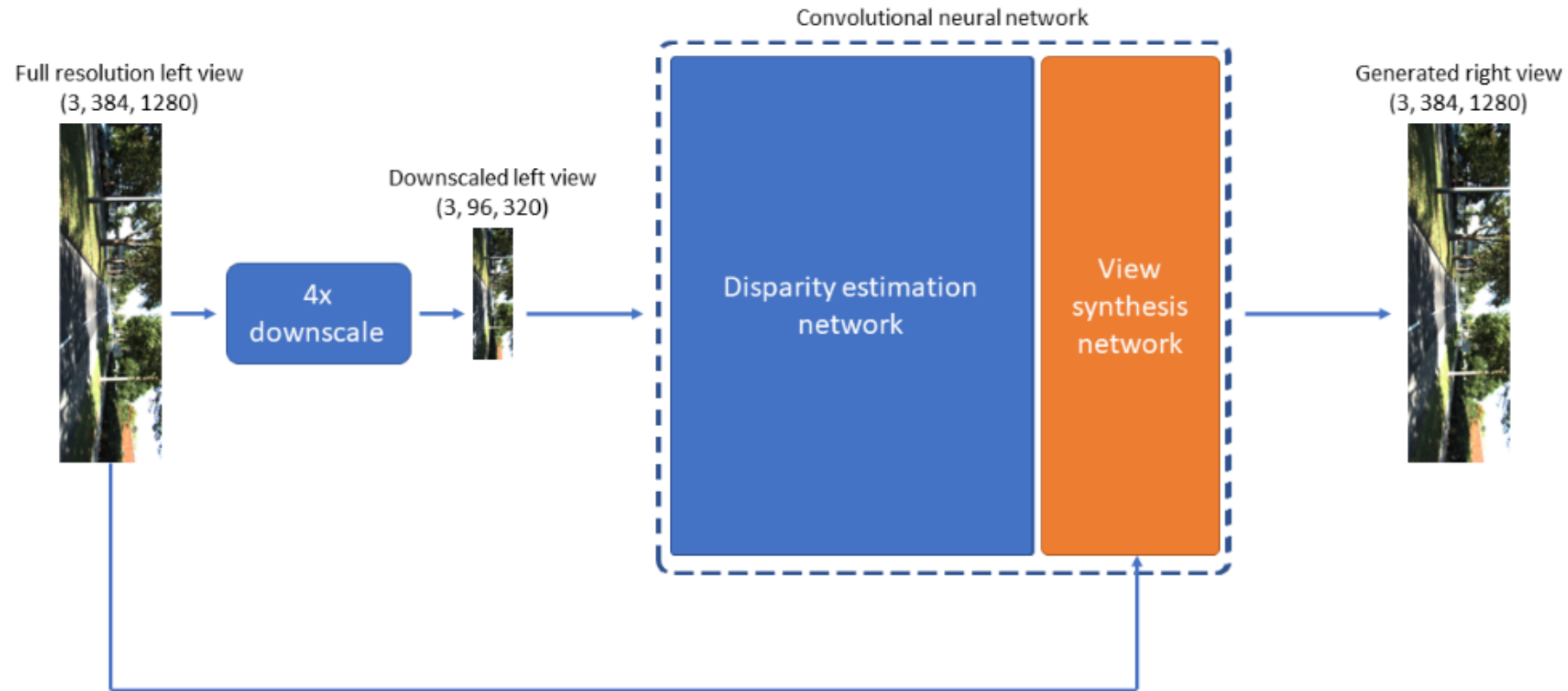
- Домени примене
  - Аутономна вожња
  - 3Д реконструкција
  - Роботика



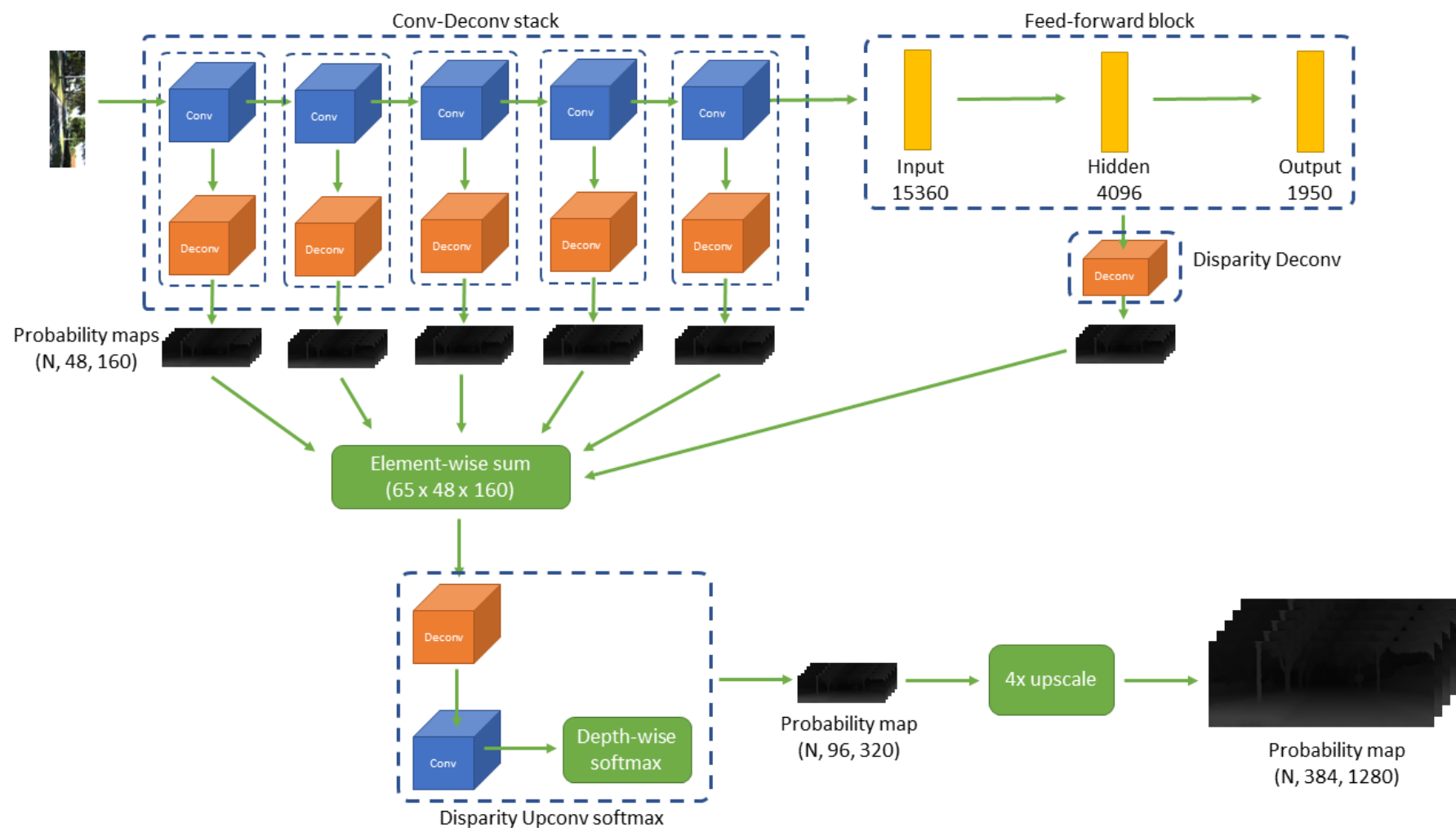
# Методологија



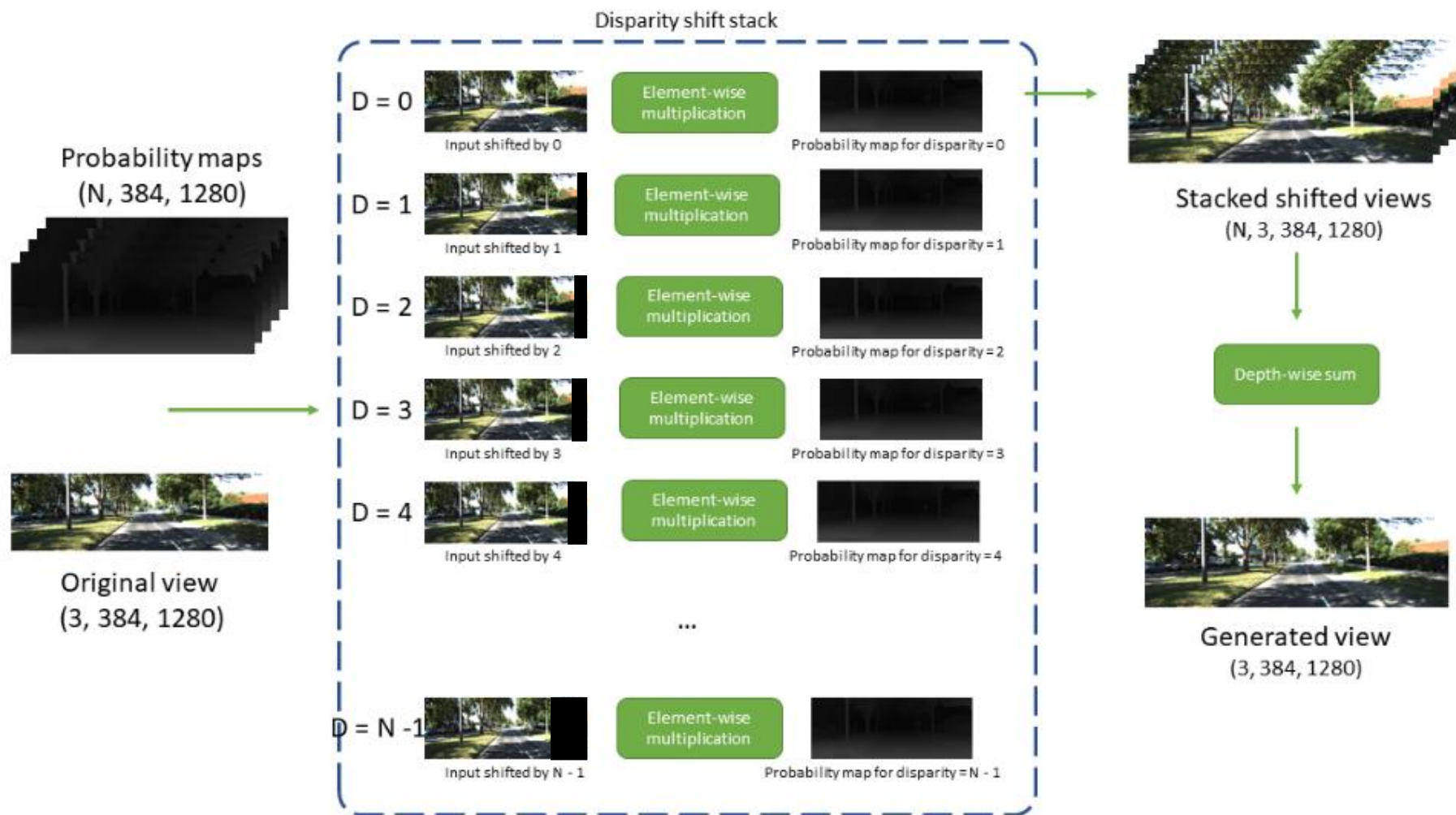
# Модул за генерисање десне референтне слике



# Модул за одређивање мапе диспаритета

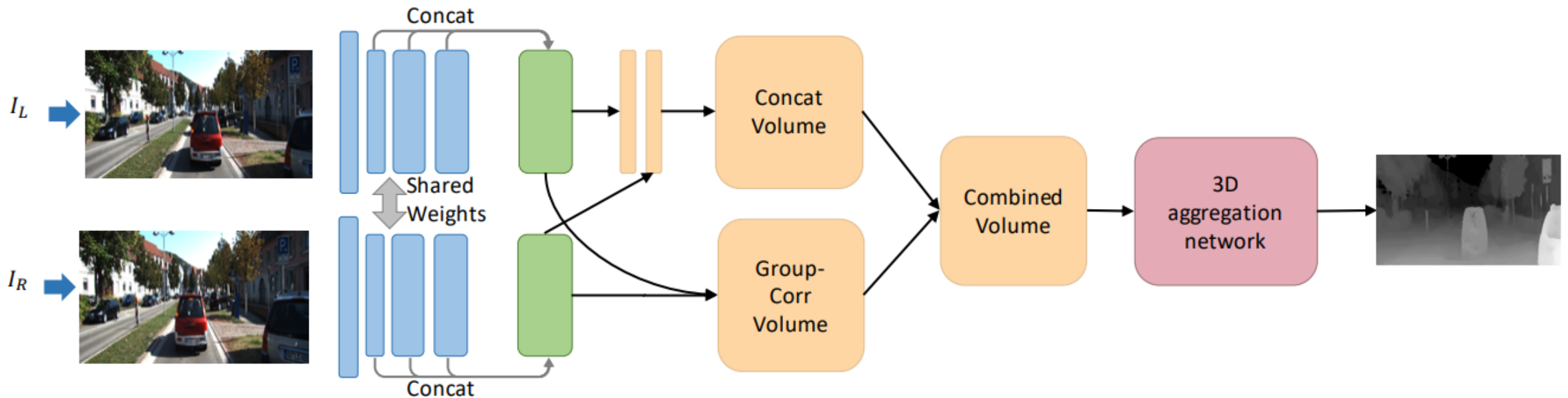


# Модул за синтезу слике



# Модули за одређивање дубине

- Стерео мечинг алгоритам – класични приступ
- *GwcNet*
- *MobileStereoNet 3D*



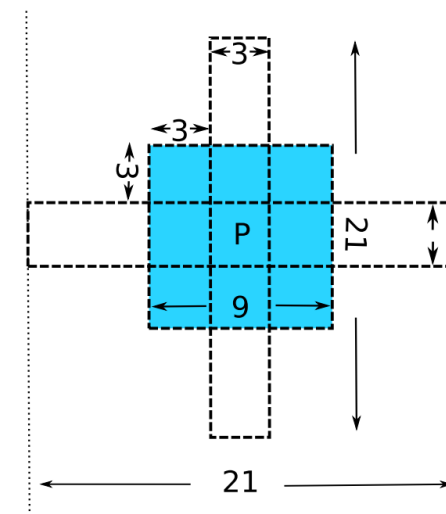
# Стерео мечинг алгоритам...

1. Конверзија улазних слика у црно-белу репрезентацију
2. Скалирање резолуције улазних слика фактором  $K$  (*MeanPooling*)
3. Конструисање *matching cost* тензора

$$C(x, y, d) = \sum_{i=-W}^W \sum_{j=-W}^W [255 - |I_{x+i, y+j}^L - I_{x+i, y+j-d}^R|]$$

4. Мулти-блок агрегације *cost* функције

$$CA(x, y, d) = \prod_{b \in B} \left[ \sum_{(i,j) \in b(x,y)} C(i, j, d) \right]$$



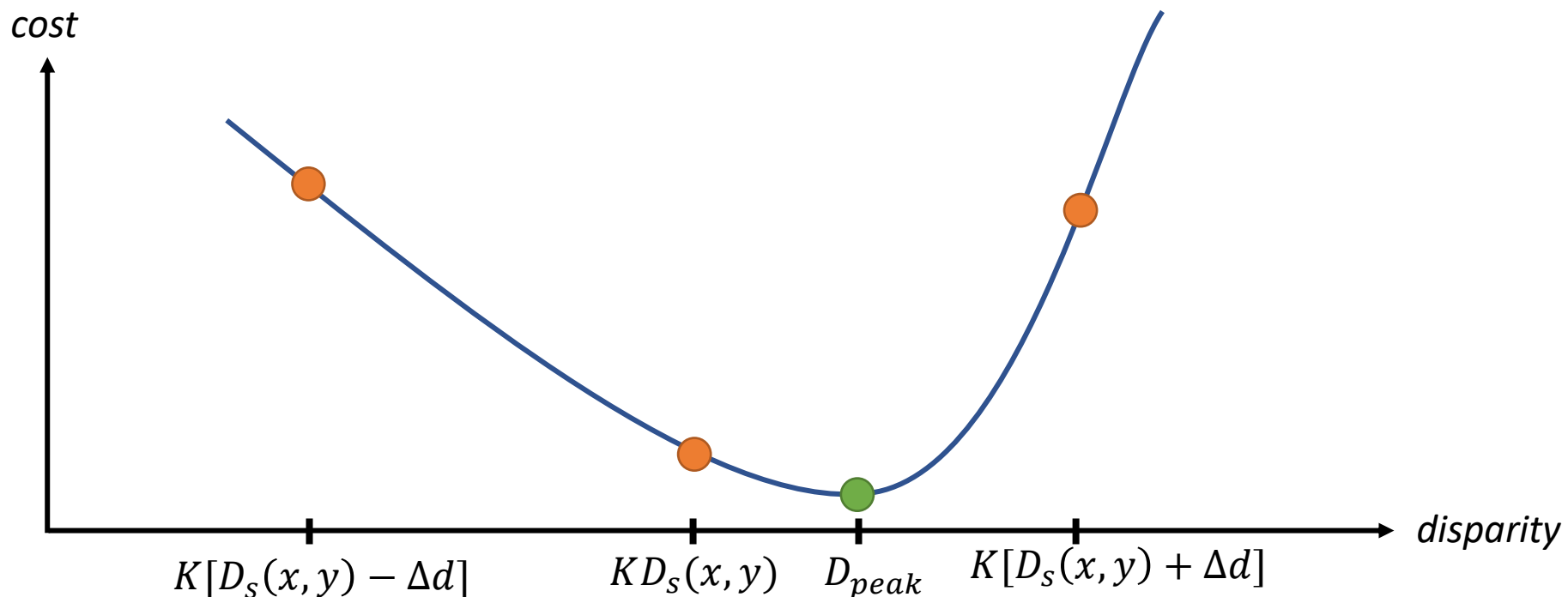


# ...Стерео мечинг алгоритам...

5. Селекција оптималног диспаритета

$$D_S(x, y) = \operatorname{argmax}\{CA(x, y, d) \mid d \in [d_{min}, d_{max}]\}$$

6. Секундарно одређивање диспаритета



## ...Стерео мечинг алгоритам

7. Повећање резолуције мапе диспаритета

$$D(K \cdot x, K \cdot y) = K \cdot D_S(x, y)$$

8. Вертикално попуњавање недостајућих вредности диспаритета

9. Хоризонтално попуњавање недостајућих вредности диспаритета

$$|D(x, y) - D(x + K, y)| \geq T$$

није

јесте

$$D(x, y + i) = D(x, y) + i \cdot \frac{[D(x + K, y) - D(x, y)]}{K},$$

за  $0 < i < K$

$$D(x, y + i) = \begin{cases} |I_{x+i,y} - I_{x,y}| \leq |I_{x+i,y} - I_{x+K,y}|, & D(x, y) \\ |I_{x+i,y} - I_{x,y}| > |I_{x+i,y} - I_{x+K,y}| & D(x + K, y) \end{cases}$$

за  $0 < i < K$

# Скупови података

## ***KITTI Raw*** скуп података

- Резолуција (375 × 1242)
- Без обележених мапа диспаратитета
- 3000 тренинг парова
- 864 тест парова



## ***KITTI 2015*** скуп података

- Подскуп *Raw* скупа података са прецизно обележеним мапама диспаратитета
- 200 тренинг парова
- 200 тест парова



# Експерименти

- Модел за генерисање десне референтне слике → 200K итерација на *KITTI Raw* скупу података
- *GwcNet* и *MobileStereoNet 3D* дотренирани на *KITTI 2015* скупу података → 45K итерација
- Евалуација са и без коришћена модула за генерисање десне референтне слике на *KITTI Raw* скупу података

# Евалуација решења

- *Mean Absolute Error*
- *D1* мера
  - Проценат пиксела за које важи:
$$|D_{est} - D_{true}| \geq \max(3, 0.05 \cdot D_{true})$$
- Број обрађених фрејмова у секунди – ФПС
- Подаци са сензора (Лидар) → генерисање обележених мапа диспаритета (100K тачака по стерео пару, ~20%)

# Резултати и дискусија...

- Са употребом модула за генерисање десне референтне слике

Метод	<i>D1</i>	<i>MAE</i>	<i>FPS</i>
Стерео мечинг	0.345	4.296	<b>30</b>
<i>GwcNet</i>	<b>0.270</b>	<b>2.856</b>	6
<i>MobileStereoNet 3D</i>	0.301	3.434	4

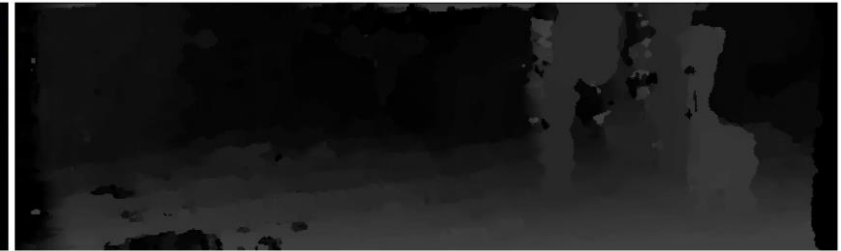
- Без употребе модула за генерисање десне референтне слике

Метод	<i>D1</i>	<i>MAE</i>	<i>FPS</i>
Стерео мечинг	0.2785	4.335	<b>91</b>
<i>GwcNet</i>	<b>0.063</b>	1.286	7
<i>MobileStereoNet 3D</i>	0.071	<b>1.269</b>	4



# ...Резултати и дискусија...

- Генерисање десне слике + Стерео мечинг алгоритам



- Генерисање десне слике + *GwcNet*



- Генерисање десне слике + *MobileStereoNet 3D*



# ...Резултати и дискусија

- Могућа побољшања:
  - Коришћење неуронских мрежа заснованих на архитектури трансформера за одређивање дубине
  - Претпроцесирање РГБ слика у циљу бољег рада са униформним регионима
  - Коришћење разређених блокова приликом рачунања агрегације *cost* функције
  - Комбиновање класичних приступа и техника машинског учења



Питања?