In [1]:
```
import sys
sys.path.append("../")
from ortho_lib import *
import pandas as pd
import os
```

In [2]:
```
data_dir = 'transformed_data'
category = 'Category_1'
patientID = 1
exercise = '/RF1'
path1 = '../transformed_data/Category_1/'
path2 = '../transformed_data/Category_2/'
path3 = '../transformed_data/Category_3/'
path4 = '../transformed_data/Category_4/'

df= pd.DataFrame()

patientID1 = os.listdir(path1)
patientID2 = os.listdir(path2)
patientID3 = os.listdir(path3)
patientID4 = os.listdir(path4)



def patients_to_df(df, patientID, path):
    print(patientID)
    for patient in patientID:
        pathex = path + patient + exercise + '.txt'
        try:
            df_patient = exercise_to_df(pathex)
        except FileNotFoundError:
            print(patient + "file not found")
        df_patient['patientID'] = patient
        df = df.append([df_patient])
    return df


df_cat1 = patients_to_df(df, patientID1, path1)
df_cat1.info()
df_cat2 = patients_to_df(df, patientID2, path2)

df_cat3 = patients_to_df(df, patientID3, path3)

df_cat4 = patients_to_df(df, patientID4, path4)

df_cat1['sensor'] = df_cat1['sensor'].astype(int)
df_cat2['sensor'] = df_cat2['sensor'].astype(int)
df_cat3['sensor'] = df_cat3['sensor'].astype(int)
df_cat4['sensor'] = df_cat4['sensor'].astype(int)
```

In [3]:
```
df_cat1 = df_cat1.set_index(['patientID', 'frame'], drop=True, inplace=False, verify_integrity=False)
#df_cat1 = df_cat1.drop(columns=['z', 'y'])

df_cat2 = df_cat2.set_index(['patientID', 'frame'], drop=True, inplace=False, verify_integrity=False)
#df_cat2 = df_cat2.drop(columns=['z', 'y'])


df_cat3 = df_cat3.set_index(['patientID', 'frame'], drop=True, inplace=False, verify_integrity=False)
#df_cat3 = df_cat3.drop(columns=['z', 'y'])


df_cat4 = df_cat4.set_index(['patientID', 'frame'], drop=True, inplace=False, verify_integrity=False)
#df_cat4 = df_cat4.drop(columns=['z', 'y'])


df_cat4
```

Out[3]:

		x	y	z	sensor
patientID	frame				
35	0	-1.050012	-0.677835	-1.788473	2
	0	0.000000	0.000000	0.000000	3
	0	-0.487536	0.574265	0.443750	4
	0	-0.357021	0.859713	-0.618892	5
	0	0.223138	0.904341	-1.379246	6
...	...	...	...	...	...
24	34	-0.442025	0.858565	-0.761021	5
	34	0.101938	1.021138	-1.518526	6
	34	-0.551872	-0.601083	0.378753	7
	34	-0.458721	-0.900746	-0.703900	8
	34	0.027242	-1.116925	-1.486011	9

11504 rows × 4 columns

Research difference x

Movement x sensor 5

In [5]:
```
df_cat1_diff_x_5 = df_cat1[df_cat1['sensor'] == 5]
df_cat2_diff_x_5 = df_cat2[df_cat2['sensor'] == 5]
df_cat3_diff_x_5 = df_cat3[df_cat3['sensor'] == 5]
df_cat4_diff_x_5 = df_cat4[df_cat4['sensor'] == 5]


df_cat1_diff_x_5 = df_cat1_diff_x_5.groupby('patientID')['x'].agg(['min', 'max'])
df_cat2_diff_x_5 = df_cat2_diff_x_5.groupby('patientID')['x'].agg(['min', 'max'])
df_cat3_diff_x_5 = df_cat3_diff_x_5.groupby('patientID')['x'].agg(['min', 'max'])
df_cat4_diff_x_5 = df_cat4_diff_x_5.groupby('patientID')['x'].agg(['min', 'max'])

diff1 = abs(df_cat1_diff_x_5['min'] - df_cat1_diff_x_5['max'])
diff2 = abs(df_cat2_diff_x_5['min'] - df_cat2_diff_x_5['max'])
diff3 = abs(df_cat3_diff_x_5['min'] - df_cat3_diff_x_5['max'])
diff4 = abs(df_cat4_diff_x_5['min'] - df_cat4_diff_x_5['max'])


df_cat1_diff_x_5['diff'] = diff1
df_cat2_diff_x_5['diff'] = diff2
df_cat3_diff_x_5['diff'] = diff3
df_cat4_diff_x_5['diff'] = diff4

print(df_cat1_diff_x_5['diff'].mean())
print(df_cat2_diff_x_5['diff'].mean())
print(df_cat3_diff_x_5['diff'].mean())
print(df_cat4_diff_x_5['diff'].mean())
```

```
0.6990775519397531
0.7721472225136001
0.7527578165630289
0.589470469292859
```

Movement x sensor 8

In [6]:
```


df_cat1_diff_x_8 = df_cat1[df_cat1['sensor'] == 8]
df_cat2_diff_x_8 = df_cat2[df_cat2['sensor'] == 8]
df_cat3_diff_x_8 = df_cat3[df_cat3['sensor'] == 8]
df_cat4_diff_x_8 = df_cat4[df_cat4['sensor'] == 8]




df_cat1_diff_x_8 = df_cat1_diff_x_8.groupby('patientID')['x'].agg(['min', 'max'])
df_cat2_diff_x_8 = df_cat2_diff_x_8.groupby('patientID')['x'].agg(['min', 'max'])
df_cat3_diff_x_8 = df_cat3_diff_x_8.groupby('patientID')['x'].agg(['min', 'max'])
df_cat4_diff_x_8 = df_cat4_diff_x_8.groupby('patientID')['x'].agg(['min', 'max'])

diff1 = abs(df_cat1_diff_x_8['min'] - df_cat1_diff_x_8['max'])
diff2 = abs(df_cat2_diff_x_8['min'] - df_cat2_diff_x_8['max'])
diff3 = abs(df_cat3_diff_x_8['min'] - df_cat3_diff_x_8['max'])
diff4 = abs(df_cat4_diff_x_8['min'] - df_cat4_diff_x_8['max'])


df_cat1_diff_x_8['diff'] = diff1
df_cat2_diff_x_8['diff'] = diff2
df_cat3_diff_x_8['diff'] = diff3
df_cat4_diff_x_8['diff'] = diff4

print(df_cat1_diff_x_8['diff'].mean())
print(df_cat2_diff_x_8['diff'].mean())
print(df_cat3_diff_x_8['diff'].mean())
print(df_cat4_diff_x_8['diff'].mean())
```

```
0.7079527984085348
0.7756107301838007
0.6986572249854184
0.617614593934907
```

Research difference Z

Movement difference sensor 5

In [7]:
```
df_cat1_diff_z_5 = df_cat1[df_cat1['sensor'] == 5]
df_cat2_diff_z_5 = df_cat2[df_cat2['sensor'] == 5]
df_cat3_diff_z_5 = df_cat3[df_cat3['sensor'] == 5]
df_cat4_diff_z_5 = df_cat4[df_cat4['sensor'] == 5]


df_cat1_diff_z_5 = df_cat1_diff_z_5.groupby('patientID')['z'].agg(['min', 'max'])
df_cat2_diff_z_5 = df_cat2_diff_z_5.groupby('patientID')['z'].agg(['min', 'max'])
df_cat3_diff_z_5 = df_cat3_diff_z_5.groupby('patientID')['z'].agg(['min', 'max'])
df_cat4_diff_z_5 = df_cat4_diff_z_5.groupby('patientID')['z'].agg(['min', 'max'])

diff1 = abs(df_cat1_diff_z_5['min'] - df_cat1_diff_z_5['max'])
diff2 = abs(df_cat2_diff_z_5['min'] - df_cat2_diff_z_5['max'])
diff3 = abs(df_cat3_diff_z_5['min'] - df_cat3_diff_z_5['max'])
diff4 = abs(df_cat4_diff_z_5['min'] - df_cat4_diff_z_5['max'])


df_cat1_diff_z_5['diff'] = diff1
df_cat2_diff_z_5['diff'] = diff2
df_cat3_diff_z_5['diff'] = diff3
df_cat4_diff_z_5['diff'] = diff4

print(df_cat1_diff_z_5['diff'].mean())
print(df_cat2_diff_z_5['diff'].mean())
print(df_cat3_diff_z_5['diff'].mean())
print(df_cat4_diff_z_5['diff'].mean())
```

```
1.0440223260451196
0.7727544823604835
0.8535202414134893
0.7139128888531565
```

Movement difference sensor 8

In [8]:
```
df_cat1_diff_z_8 = df_cat1[df_cat1['sensor'] == 8]
df_cat2_diff_z_8 = df_cat2[df_cat2['sensor'] == 8]
df_cat3_diff_z_8 = df_cat3[df_cat3['sensor'] == 8]
df_cat4_diff_z_8 = df_cat4[df_cat4['sensor'] == 8]

df_cat1_diff_z_8 = df_cat1_diff_z_8.groupby('patientID')['z'].agg(['min', 'max'])
df_cat2_diff_z_8 = df_cat2_diff_z_8.groupby('patientID')['z'].agg(['min', 'max'])
df_cat3_diff_z_8 = df_cat3_diff_z_8.groupby('patientID')['z'].agg(['min', 'max'])
df_cat4_diff_z_8 = df_cat4_diff_z_8.groupby('patientID')['z'].agg(['min', 'max'])

diff1 = abs(df_cat1_diff_z_8['min'] - df_cat1_diff_z_8['max'])
diff2 = abs(df_cat2_diff_z_8['min'] - df_cat2_diff_z_8['max'])
diff3 = abs(df_cat3_diff_z_8['min'] - df_cat3_diff_z_8['max'])
diff4 = abs(df_cat4_diff_z_8['min'] - df_cat4_diff_z_8['max'])


df_cat1_diff_z_8['diff'] = diff1
df_cat2_diff_z_8['diff'] = diff2
df_cat3_diff_z_8['diff'] = diff3
df_cat4_diff_z_8['diff'] = diff4

print(df_cat1_diff_z_8['diff'].mean())
print(df_cat2_diff_z_8['diff'].mean())
print(df_cat3_diff_z_8['diff'].mean())
print(df_cat4_diff_z_8['diff'].mean())
```

```
1.0392412967910265
0.7469087123840794
0.8887016827181824
0.7616880743592568
```

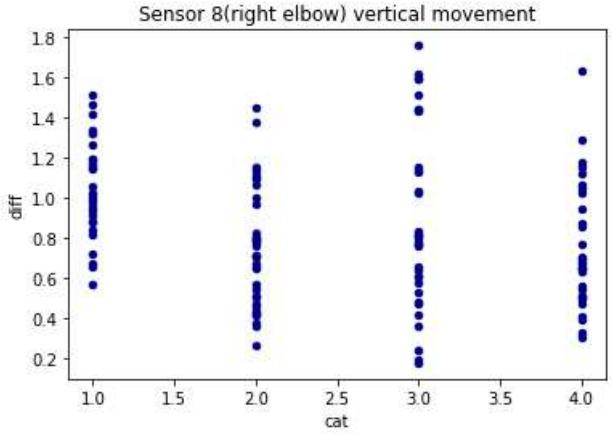
Plot difference combined DF per sensor

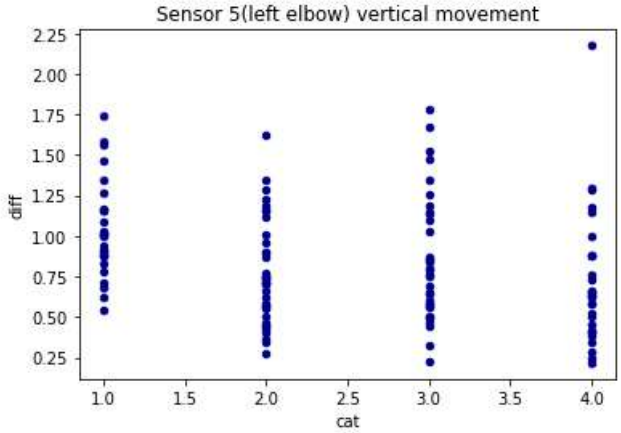
In [9]:
```
df_cat1_diff_z_8['cat'] = 1
df_cat2_diff_z_8['cat'] = 2
df_cat3_diff_z_8['cat'] = 3
df_cat4_diff_z_8['cat'] = 4

df_cat1_diff_z_5['cat'] = 1
df_cat2_diff_z_5['cat'] = 2
df_cat3_diff_z_5['cat'] = 3
df_cat4_diff_z_5['cat'] = 4

df_cat1_diff_x_8['cat'] = 1
df_cat2_diff_x_8['cat'] = 2
df_cat3_diff_x_8['cat'] = 3
df_cat4_diff_x_8['cat'] = 4

df_cat1_diff_x_5['cat'] = 1
df_cat2_diff_x_5['cat'] = 2
df_cat3_diff_x_5['cat'] = 3
df_cat4_diff_x_5['cat'] = 4

df_combined_diff_z_8 = df_cat1_diff_z_8.append([df_cat2_diff_z_8, df_cat3_diff_z_8, df_cat4_diff_z_8])

df_combined_diff_z_5 = df_cat1_diff_z_5.append([df_cat2_diff_z_5, df_cat3_diff_z_5, df_cat4_diff_z_5])

df_combined_diff_x_5 = df_cat1_diff_x_5.append([df_cat2_diff_x_5, df_cat3_diff_x_5, df_cat4_diff_x_5])

df_combined_diff_x_8 = df_cat1_diff_x_8.append([df_cat2_diff_x_8, df_cat3_diff_x_8, df_cat4_diff_x_8])

df_combined_diff_z_8.plot.scatter(x='cat', y='diff', ylabel= 'difference between min and max', c='DarkBlue', title='Sensor 8(righ

df_combined_diff_z_5.plot.scatter(x='cat', y='diff', c='DarkBlue', title='Sensor 5(left elbow) vertical movement')

df_combined_diff_x_8.plot.scatter(x='cat', y='diff', c='DarkBlue', title='Sensor 8(right elbow) horizontal movement')

df_combined_diff_x_5.plot.scatter(x='cat', y='diff', c='DarkBlue', title='Sensor 5(left elbow) horizontal movement')

df_combined_diff_x_5
```

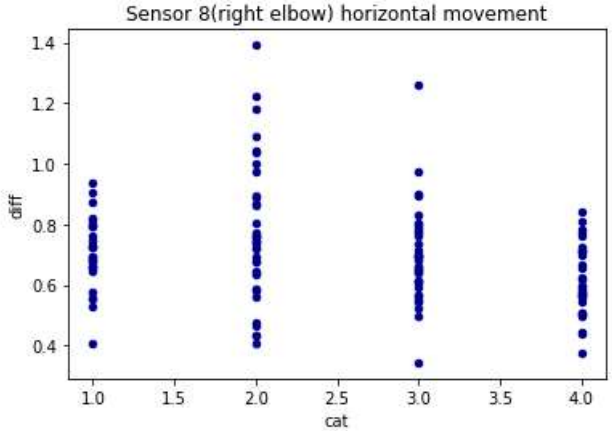
Out[9]:

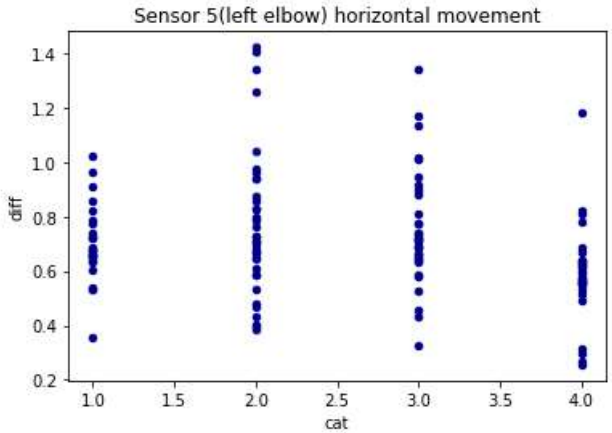
patientID	min	max	diff	cat
1	-1.028998	-0.341758	0.687240	1
10	-0.975701	-0.316814	0.658887	1
11	-1.156232	-0.130857	1.025375	1
12	-1.093970	-0.461833	0.632137	1
13	-1.043996	-0.323687	0.720309	1
...	...	...	...	...
5	-0.889684	-0.633747	0.255936	4
6	-1.041541	-0.403237	0.638304	4
7	-1.153025	-0.514139	0.638886	4
8	-0.822327	-0.524074	0.298253	4
9	-1.229140	-0.699185	0.529955	4

127 rows × 4 columns


Sensor 8(right elbow) vertical movement
diff
cat


Sensor 5(left elbow) vertical movement
diff
cat


Sensor 8(right elbow) horizontal movement
diff
cat


Sensor 5(left elbow) horizontal movement
diff
cat

Klein model proberen te maken voor testen.

In [10]:
```
#df_test = df_combined_diff_z_8[abs(df_combined_diff_z_8['diff']-df_combined_diff_z_8['diff'].mean())
 #                   <= (3*df_combined_diff_z_8['diff'].std())]


#q_low = df_combined_diff_z_8["diff"].quantile(0.10)
#q_hi  = df_combined_diff_z_8["diff"].quantile(0.95)


#df_test = df_combined_diff_z_8[(df_combined_diff_z_8['diff'] < q_hi) & (df_combined_diff_z_8["diff"] > q_low) ]

df_test = df_combined_diff_x_8

df_test = df_test[df_test['cat'] != 3]
df_test = df_test[df_test['cat'] != 2]
```

In [11]:
```
df_test.plot.scatter(x='cat', y='diff', c='DarkBlue', title='Sensor 8(left elbow) horizontal movement')


df_test
```

```
from sklearn.model_selection import train_test_split
from sklearn.model_selection import StratifiedKFold
import numpy as np

#splitten test en train set

X = np.asarray(df_test[['diff']])
y = np.asarray(df_test[['cat']])
skf = StratifiedKFold()

for train_index, test_index in skf.split(X, y):
    print("TRAIN:", train_index, "TEST:", test_index)
    X_train, X_test = X[train_index], X[test_index]
    y_train, y_test = y[train_index], y[test_index]



#x_train, x_test, y_train, y_test = train_test_split(X, y,random_state=2, test_size=0.2)

print(X_train.shape, y_train.shape, X_test.shape, y_test.shape)
```

```
TRAIN: [ 6  7  8  9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 33 34 35
 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56] TEST: [ 0  1  2  3  4  5 27 28 29 30 31 32]
TRAIN: [ 0  1  2  3  4  5 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29
 30 31 32 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56] TEST: [ 6  7  8  9 10 11 33 34 35 36 37 38]
TRAIN: [ 0  1  2  3  4  5  6  7  8  9 10 11 17 18 19 20 21 22 23 24 25 26 27 28
 29 30 31 32 33 34 35 36 37 38 45 46 47 48 49 50 51 52 53 54 55 56] TEST: [12 13 14 15 16 39 40 41 42 43 44]
TRAIN: [ 0  1  2  3  4  5  6  7  8  9 10 11 12 13 14 15 16 22 23 24 25 26 27 28
 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 51 52 53 54 55 56] TEST: [17 18 19 20 21 45 46 47 48 49 50]
TRAIN: [ 0  1  2  3  4  5  6  7  8  9 10 11 12 13 14 15 16 17 18 19 20 21 27 28
 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50] TEST: [22 23 24 25 26 51 52 53 54 55 56]
(46, 1) (46, 1) (11, 1) (11, 1)
```

In [13]:
```
from sklearn.linear_model import LogisticRegression

#make instance of model



logistic_reg = LogisticRegression(solver='liblinear')

logistic_reg.fit(X_train,y_train.ravel())
```

Out[13]:
```
LogisticRegression(solver='liblinear')
```

In [14]:
```
y_predict = logistic_reg.predict(X_test)

print(y_predict)
```

```
[1 4 4 1 1 4 4 4 1 4 4]
```

In [15]:
```
score = logistic_reg.score(X_test, y_test)
print(score)
```

```
0.7272727272727273
```